Methods for efficient drug development in neuropsychiatric diseases

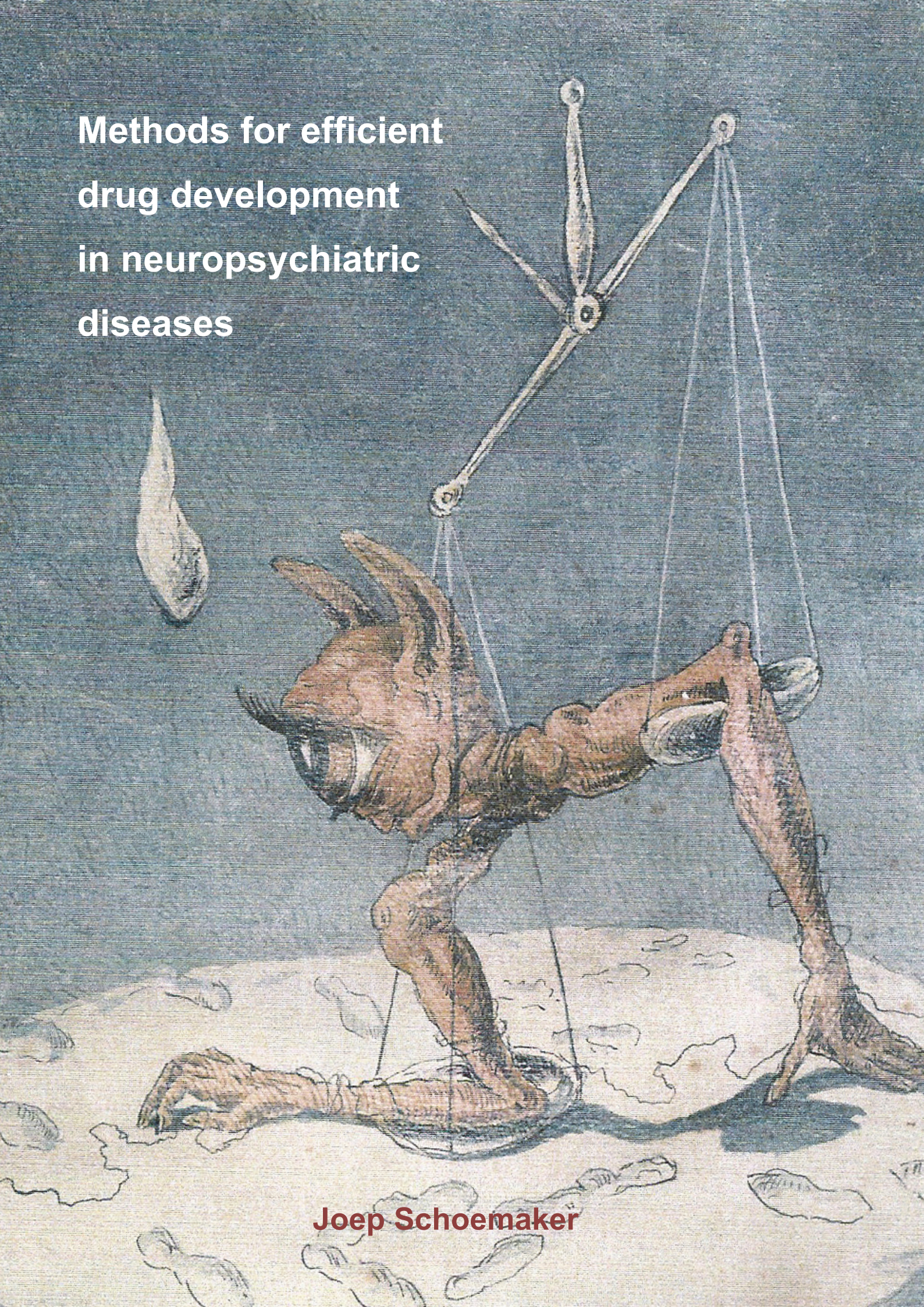
Schoemaker, Joep

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Schoemaker, J. (2018). *Methods for efficient drug development in neuropsychiatric diseases*. GVO drukkers & vormgevers B.V. | Ponsen & Looijen.

Methods for efficient drug development in neuropsychiatric diseases



Joep Schoemaker

Methods

for efficient

drug development

in neuropsychiatric diseases

Joep Schoemaker

Cover illustration: ‘Подозретильность’ (Susceptibility) — Michail Paule (189?-1939)
Painted during his stay at Saratov State Medical University Hospital (1930-1937)

ISBN 978-94-6332-384-0
NUR-code 883 – Medical & socio-medical sciences

Printed by GVO drukkers & vormgevers B.V. | Ponsen & Looijen, Ede

Methods for efficient drug development in neuropsychiatric diseases

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan Tilburg University
op gezag van de rector magnificus,
prof.dr. E.H.L. Aarts,

in het openbaar te verdedigen ten overstaan van een
door het college voor promoties aangewezen commissie
in de aula van de Universiteit

op woensdag 10 oktober 2018 om 14.00 uur

door

Josephus Hubertus Schoemaker,
geboren te Vught.

Promotores:

Prof. dr. A.J.J.M. Vingerhoets

Prof. dr. R.A. Emsley

Promotiecommissie:

Prof. dr. M. De Hert

Prof. dr. W.J. Kop

Prof. dr. A.J.M. Loonen

Prof.em.dr. J.M.A. Sitsen

Prof. dr. J.A. Swinkels

CONTENTS

Chapter 1

Background

General introduction

Outline of the thesis

GENERAL INTRODUCTION

Drug discovery and clinical development is a risky, costly and time-consuming enterprise and traditionally occurs in stages, although the clinical phases of drug testing usually show substantial overlap (Figure 1.1).

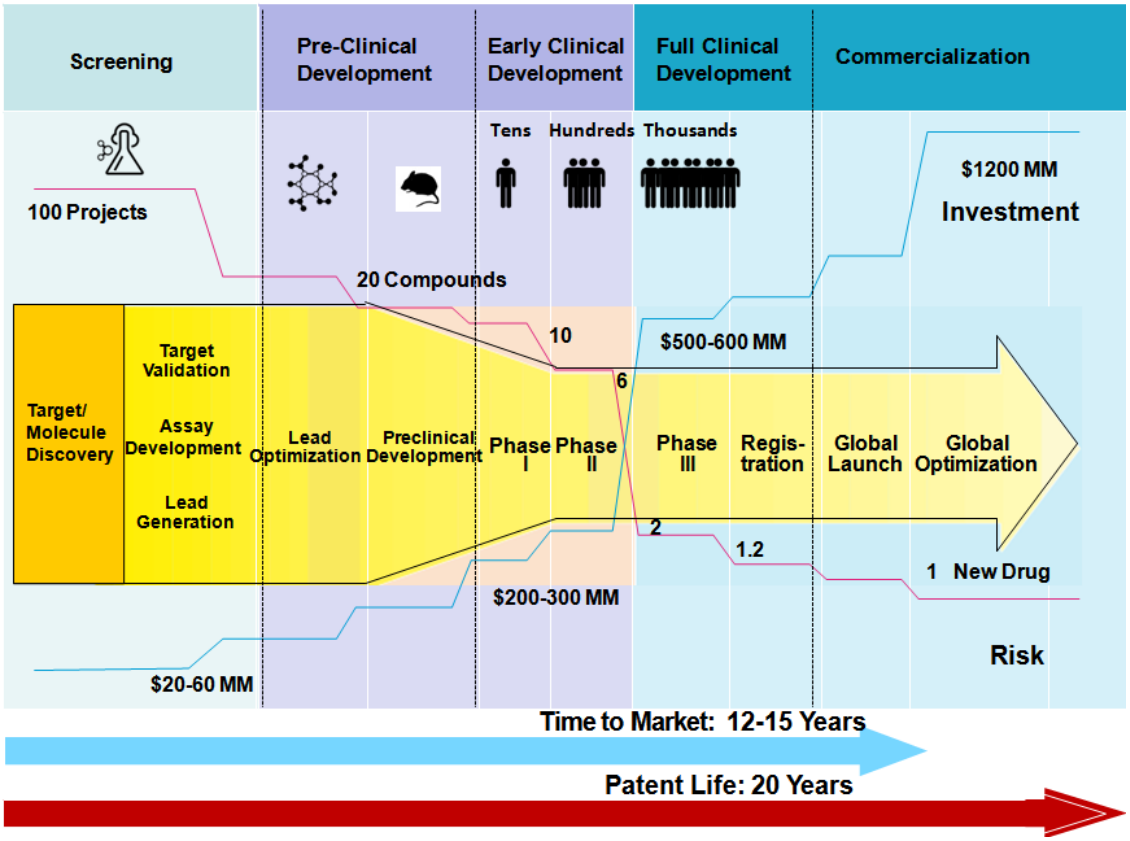


Figure 1.1

Drug discovery and development process, showing the gradual decline of compounds in development against growing costs, whereby any delay in development time threatens the profitability of a new drug because of limited patent duration.[1-3]

Over the last five decades, pharmaceutical companies have endeavored to develop new drugs targeting psychiatric diseases with varying rates of success. In some areas, substantial progress has been made with pharmacological treatment (such as in anxiety and insomnia) whereas particular diseases remain notoriously difficult to treat. Novel antidepressants (including the widely used serotonin reuptake inhibitors) have been described as being only slightly more effective than placebo in the treatment of major depressive disorder (MDD), and not more efficacious than drugs such as amitriptyline, developed more than 40 years ago.[4,5]

The same holds for second-generation antipsychotics (SGA), which may be slightly better tolerated in some respect, but still may not show better efficacy than first-generation antipsychotics (FGA) such as perphenazine.[6,7] Despite a vastly growing number of drugs approved, many patients with schizophrenia or depression fail to respond to treatment, or keep suffering from residual symptoms affecting their daily life. In a survey among almost 6000 MDD patients, 30.9% considered themselves to have responded to treatment, 31.2% considered themselves to have only partially responded, whereas 37.9% did not consider themselves to have responded at all.[8] In a large, US National Institutes of Health trial sponsored trial, known as CATIE, 74% of patients with schizophrenia discontinued drug use within 18 months therapy due to either poor tolerability or insufficient therapeutic effect.[6] A meta-analysis, based on 38 randomized, controlled trials (RCTs) in schizophrenia with 7,323 participants, demonstrated a relatively small absolute gain of 17% in response rate to SGA compared to placebo (overall 41% versus 24% responded to treatment with SGA drugs and placebo, respectively).[9] In a comprehensive analysis of 167 RCTs involving 28,102 participants with mainly chronic schizophrenia, approximately twice as many patients improved with antipsychotics as with placebo, but only a minority (24% on drug versus 14% on placebo) showed 'good response'.[7]

Challenges in psychiatric drug development

The limited response to drug treatment in psychiatric diseases may be attributed to a variety of factors. Diagnosis is typically by observation, detection is late, and prediction is poor. This is because the neurobiology of most disorders is complex, not confined to a single neuroanatomical circuit, and their etiology in many cases largely unknown.[10] Although people with a specific genetic background may show increased vulnerability to develop a specific mental disorder, validated biomarkers for disease severity are genuinely lacking, and for most syndromes, there are no objective criteria to base a diagnosis upon.[11] As a result, diagnostic boundaries are rather arbitrary and may not reflect distinct neurobiological entities, so that symptoms from recognizable syndromes may show substantial overlap.[12] Individual symptoms (such as depressed mood and hallucinations) can be mediated and modulated by a variety of neurobiological mechanisms, and – although receptor binding data from imaging studies may aid in the demonstration of a molecule's target engagement - hardly any dose-linear relationships have been established between neurotransmitter receptor occupancies and

clinical response to drug treatment.[13] Last but not least, drug treatment effects are quantified using standardized psychometric assessments that are based on clinical interviews, observations, self-reflections and/or cognitive testing, each with their respective limitations and risks for bias.

Inconclusive clinical trials

As long as the chance for clinical improvement remains limited, patients will be difficult to recruit for RCTs and keep motivated staying compliant with prescribed treatment regimens under experimental settings. Due to the challenges described above, the timelines (and costs) for developing new drugs targeting the central nervous system (CNS) are substantially higher than for most other therapeutic classes (Figure 1.2).[3,14]

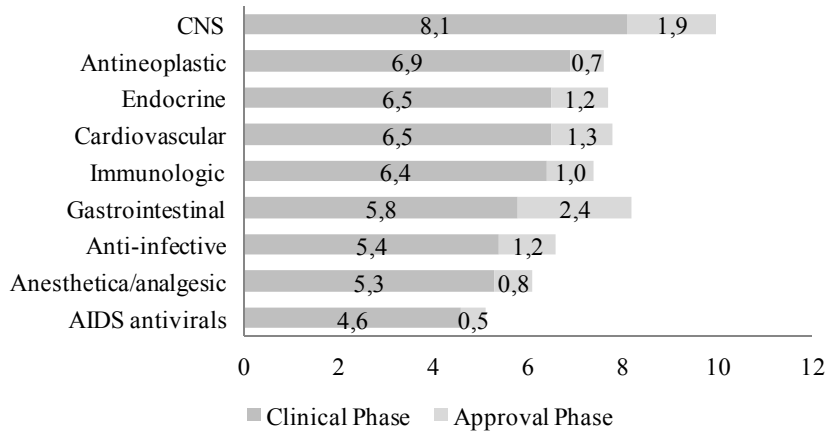


Figure 1.2

Mean clinical and approval phase times (in years) for new drugs, registered between 2005 and 2009, grouped by therapeutic class.[3]

It is therefore not surprising that many pharmaceutical companies (including Novartis, Glaxo-Smith-Kline, Astra Zeneca, Sanofi Aventis, and Merck) have gradually abandoned their efforts to develop new drugs for neuropsychiatric diseases, because not only the costs, but also risks for failure become more and more considered to be unacceptably high (Figures 1.3 & 1.4).[2,14,15] An additional, major complication is caused by the increased response rate to placebo observed over the last ten years in RCTs in psychiatric diseases, including MDD, schizophrenia, and bipolar mania (see Chapter 2), resulting in negative and failed clinical

trials.* This has resulted in an increased interest in methods to enhance the success rate or signal detection in clinical trials.

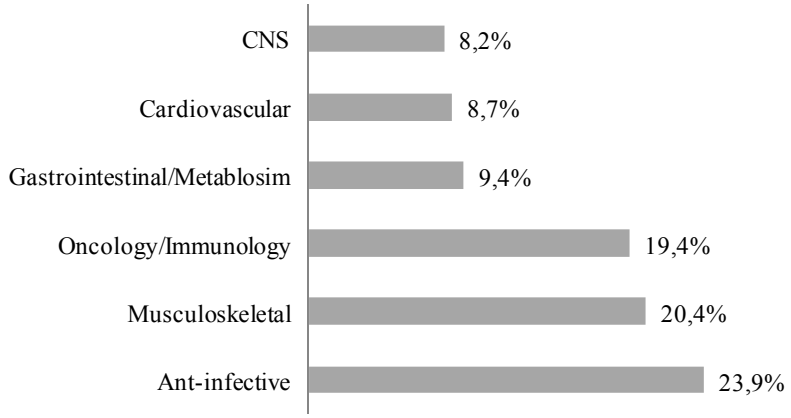


Figure 1.3

Clinical approval success rate of investigational drugs in CNS and other therapeutic areas.[2]

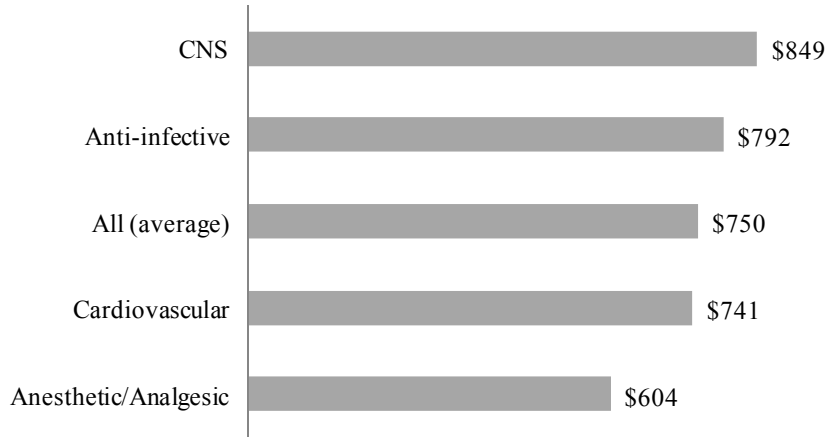


Figure 1.4

Capitalized clinical development costs (in millions) by therapeutic area.[2]

* A trial is called negative, when an active drug- control differs significantly from placebo and the investigational drug does not, and failed, when neither drug differs significantly from placebo.

OUTLINE OF THE THESIS

Tested drug development strategies

In order to retain costs and efforts, several strategies can be adopted to develop new drugs for neuropsychiatric diseases during the various phases of development. Some of these are described below and were tested in multicentre studies (Chapters 3-5 of the thesis).

In the early development phase, *signal enhancement strategies* may be implemented to mitigate placebo response and increase the accuracy of outcome measurements. These may include adaptive protocol designs, use of specific outcome measures, and exclusion of patients with an increased likelihood of responding to placebo. The continuous search for mediators of placebo response in depression studies is reviewed in **Chapter 2**.

As long as a novel mechanism of action needs to be proven, it is prudent to test new drugs as *adjunctive treatment* to established drugs, targeting unmet needs in partially responsive patients. Such an approach was adopted for the study presented in **Chapter 3**, which also involved novel methods to enhance rater accuracy and statistical techniques to discriminate direct from secondary effects on the primary outcome variable.

It is not uncommon these days to enroll (already in the pre-registration phase) a relatively large sample of patients for a face-to-face comparison with a market leading product, whereby a wide range of comorbid conditions and adjunctive medications are allowed. An example of such a Phase III trial under more *'naturalistic' conditions* is provided in **Chapter 4**. An approach like this, not only facilitates the generalizability of study results to the average patient population but may also show the newcomer's benefit/risk ratio in comparison with established drugs.

Continued access to the experimental drug upon successful trial completion can be a very important asset for patients volunteering to participate in an RCT. This can be accomplished through a so-called *'humanitarian' extension* study, allowing the sponsor at the same time to collect long-term outcome data. The study presented in **Chapter 5** can be regarded as a successful example in this respect.

Normally speaking, a well designed clinical program is required to demonstrate the effectiveness of a test drug. Important to keep in mind when interpreting the results of RCTs is that the absence of evidence does not automatically imply evidence of absence.[16] Response to placebo, inadequate dosages tested, and missing data (e.g., due to drop-out), but in particular not selecting the right patients, dosing, and endpoint may be common reasons for

the occurrence of so-called inconclusive trials (i.e., a test drug not showing superiority over placebo, despite potentially exhibiting biological activity).[17] Due to patients' lack of insight in their own disease, disturbing side effects of treatment, cognitive impairment, social isolation, common symptoms inherent to the disease itself (such as suspiciousness), and comorbid substance abuse, the challenges of drug treatment and prevalence of nonadherence are relatively high among patients with schizophrenia.[18,19] By rule of thumb, investigators will need to establish a good relationship with their patients before being able to enroll them successfully as participants in an RCT.[20,21] More than in other indications, *withdrawal of consent* is a major reason for drop-out in schizophrenia studies. Since a high drop-out rate decreases the power of an RCT, it should be avoided at all times. **Chapter 6** explores what makes a patient decide to discontinue participation in a study.

Once a drug has been approved for prescription in a particular region, a sponsor may be able to expand its registration status through a so-called *bridging study*, whereby the effects of parallel treatment in populations from different geographic regions are compared. The objectives of such a study are (*1*) to show that the drug is effective in the new region, and (*2*) to compare the results of the study between the regions with the intent of establishing that the efficacy & safety profile of the drug is not sensitive to ethnic factors. The ultimate aim is then to gain access to a wider market without the need to repeat the full development program in the new region. An example of such a study is provided in **Chapter 7**.

Concluding remarks

Although the development strategies highlighted in this thesis are far from exhaustive, in **Chapter 8**, some concluding remarks are made with regards to the particular perspectives each may entail. Implications for further research are provided, stressing the importance of the engagement and involvement of representative patient populations in drug development trajectories, whereby a patient's motivation to participate in clinical studies is better understood and optimally exploited.

REFERENCES

1. Doroshow JH. Re-engineering early phase cancer drug development: decreasing the time from novel target to novel therapeutic. *16th Annual Drug Discovery Symposium, Chicago* 2011.
2. Miller G. Is pharma running out of brainy ideas? *Science* 2010;329(5991):502-504.
3. Kaitin KI, DiMasi JA. Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000–2009. *Clinical Pharmacology & Therapeutics* 2011;89(2):183-188.
4. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 2008;358(3):252-260.
5. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 2008;5(2):0260-0268.
6. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RS, Davis SM, Davis CE, Lebowitz BD. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine* 2005;353(12):1209-1223.
7. Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B, Samara M, Rabaioli M, Bächer S, Cipriani A. Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, Bayesian meta-analysis, and meta-regression of efficacy predictors. *American Journal of Psychiatry* 2017;174(10):927-942.
8. Knoth RL, Bolge SC, Kim E, Tran Q-V. Effect of inadequate response to treatment in patients with depression. *American Journal of Managed Care* 2010;16(8):e188-e196.
9. Leucht S, Arbter D, Engel RR, Kissling W, Davis JM. How effective are second-generation antipsychotic drugs? A meta-analysis of placebo-controlled trials. *Molecular Psychiatry* 2009;14(4):429-447.
10. Goodkind M, Eickhoff SB, Oathes DJ, Jiang Y, Chang A, Jones-Hagata LB, Ortega BN, Zaiko YV, Roach EL, Korgaonkar MS, Grieve SM, Galatzer-Levy I, Fox PT, Etkin A. Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry* 2015;72(4):305-315.
11. Insel TR, Wang PS. Rethinking mental illness. *JAMA* 2010;303(19):1970-1971.
12. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry* 2003.
13. Wong DF, Tauscher J, Grunder G. The role of imaging in proof of concept for CNS drug discovery and development. *Neuropsychopharmacology* 2009;34(1):187-203.
14. Kaitin K. *Pace of CNS drug development and FDA approvals lags other drug classes*. Boston, USA: Tufts Center for the Study of Drug Development;2012 Mar/Apr. RS 3207.
15. Johnson GS. Commercial viability of CNS drugs: Balancing the risk/reward profile. *Neurobiology of Disease* 2014;61:21-24.
16. Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence. *British Medical Journal* 1995;311(7003):485.
17. Chin R, Lee BY. *Principles and practice of clinical trial medicine.* Elsevier; 2008.
18. Cramer JA, Rosenheck R. Compliance with medication regimens for mental and physical disorders. *Psychiatric Services* 2006;49(2):196-201.
19. Valenstein M, Ganoczy D, McCarthy JF, Kim HM, Lee TA, Blow FC. Antipsychotic adherence over time among patients receiving treatment for schizophrenia: a retrospective review. *Journal of Clinical Psychiatry* 2006;67(10):1,478-1550.
20. Day JC, Bentall RP, Roberts C, Randall F, Rogers A, Cattell D, Healy D, Rae P, Power C. Attitudes toward antipsychotic medication: the impact of clinical variables and relationships with health professionals. *Archives of General Psychiatry* 2005;62(7):717-724.
21. McCabe R, Bullenkamp J, Hansson L, Lauber C, Martinez-Leal R, Rössler W, Salize HJ, Svensson B, Torres-Gonzalez F, van den Brink R. The therapeutic relationship and adherence to antipsychotic medication in schizophrenia. *PLoS One* 2012;7(4):e36080.

Chapter 2

Evaluation of placebo response factors in depression

Towards a better understanding of factors reducing power in randomized, controlled trials (I) [1]

[1] This chapter was published as:

Schoemaker, Joep H., Kilian, Sanja, Emsley, Robin, Vingerhoets, Ad J.J.M. (2018). Factors associated with placebo response in depression trials: a systematic review of published meta-analyses (1990-2017). *Neurology, Psychiatry, and Brain Research 30, 12-21.*

DOI: http:dx.doi.org/10.1016/j.npbr.2018.04.002

ABSTRACT

Background: Placebo response is common in patients with major depressive disorder (MDD) and decreases the likelihood of demonstrating drug superiority over placebo in a randomized, controlled trial (RCT). This paper aims to review the collective evidence for particular patient characteristics and trial features being associated with placebo response in MDD.

Methods: MEDLINE/PubMed publication database and Cochrane Library were searched for meta-analyses of placebo response in MDD, published in English from January 1990 to December 2017. The evidence for factors predicting a low or high placebo response was tabulated and weighted on the basis of methods, results, and quality of supporting studies.

Results: We identified 58 papers, examining the possible association of 40 different factors with placebo response in MDD. Research methods varied considerably across articles so that our reporting remained descriptive. The evidence for any factor being associated with placebo response in MDD appeared very weak to weak.

Conclusions: Despite 25 years of pooling data from RCTs in MDD, there is no single factor for which strong evidence exists that it influences placebo response.

INTRODUCTION

Sponsors of randomized controlled trials (RCTs) have been facing an increasing response rate to placebo in major depressive disorder (MDD) and other neuropsychiatric disorders over the last decades, resulting in failed studies, delayed or abandoned projects, and steep increases in Research and Development costs.[1-5] In the previous 25 years, this has led to a multitude of pooled analyses investigating predictors of placebo response in MDD.

The evidence from pooled analyses is frequently not convincing and sometimes even contradictory for the many predictors examined.[6] This could be due to sampling bias (e.g., when factors are explored on multiple occasions for their association with placebo response under non-uniform conditions), or methodological flaws, such as the use of inappropriate statistical models, regression to the mean effects, and the use of the relative efficacy of antidepressants versus placebo as outcome variable (which can only provide indirect evidence for a factor being associated with placebo response). In order to identify moderators of placebo response, it is essential to use data from as many RCTs as possible, preferably at patient-level rather than study-level, and to correct for heterogeneity in study design when executing pooled analyses. The use of individual study participant data is ideal in any meta-analysis, in that it allows standardizing the statistical analysis of each study, obtaining summary results directly, checking the assumptions of models, examining interactions, and adjusting each patient's change score by their baseline value and other patient-level characteristics.

A limited number of meta-analyses have been reviewed by several authors (here called "meta-reviews"), with conclusions sometimes being drawn without regard to the methods applied in the meta-analyses, the existence of contradictory results reported elsewhere, or without providing criteria for weighing the level of evidence coming from these meta-analyses.[6-8] On the basis of a review of 13 meta-analyses, Rutherford and Roose concluded that there is "strong" evidence for a positive association between placebo response in depression RCTs and (*1*) a lower probability of receiving placebo or multiple active treatment arms, (*2*) the average number of study sites in a RCT, and (*3*) poor rater blinding, without providing the criteria on the basis of which they reached this conclusion.[7] After reviewing 14 meta-analyses in depression (5 original, and 9 already reviewed by Rutherford and Roose), Weimer et al. reported (*1*) lower probability of receiving placebo, (*2*) low illness severity, and

(*3*) more recent RCTs to be associated with greater placebo response.[8] Papakostas et al. reviewed 23 relevant meta-analyses (of which 12 original, not yet included in the two previous meta-reviews), and reported repetitive evidence for a positive association between placebo response and (*1*) lower probability of receiving placebo, (*2*) low illness severity, and (*3*) increased visit frequency.[6] The authors of three meta-analyses (at study-level), published between 2004 and 2010, unanimously concluded that at least a lower probability of receiving placebo is likely to inflate placebo response in depression trials.[9-11] However, the results of four more recent meta-analyses (of which three at study-level) published between 2012 and 2016, strongly indicate that there is no such effect.[12-15] The inconsistency in results shows that, even when repeated findings lead to seemingly justifiable conclusions, subsequent meta-analyses exploring the same relationship may generate conflicting results, especially when data are aggregated at study-level. It underlines the need for authors of reviews to collect data from as many sources as possible, and to preferably weigh the results of individual studies on the basis of certain quality criteria.

Aims of the study

The objective of the current work is to review the collective evidence regarding associations of placebo response with trial design and patient characteristics that have been previously explored in pooled analyses, meta-regressions, and other effect models whereby the results of RCTs in MDD are combined. Unlike previous systematic reviews of meta-analyses, the aim of the current paper is to bring together the results of all previous work and grade the strength of evidence for identified predictors on the basis of pre-defined criteria for quality, quantity, and specificity of the data underlying each analysis. This approach has several advantages. Firstly, the importance of contradictory or isolated findings, whenever occurring, can be weighed to a certain extent, which reduces the risk of drawing false or unjustified conclusions. It may also help to create a better understanding of the relatively weak predictive value of results coming from pooled analyses when these are based on aggregated data at study-level rather than at patient-level. This is important particularly when only RCTs are analyzed with relatively low placebo response (i.e., excluding negative, or failed studies which were never published), or when heterogeneity in the design of underlying RCTs is not well accounted for. Finally, it may contribute to a better trial design for the demonstration of efficacy of new products in depression.

METHODS

The MEDLINE/PubMed publication database and the Cochrane Library were searched for meta-analyses and pooled-analyses (from here, all called 'meta-analyses' for the sake of simplicity) of placebo response in MDD. The search term 'placebo' was cross-referenced with the terms 'depression' or 'antidepressant,' 'response' or 'effect,' and 'trial' in Title/Abstracts to identify articles focusing on contributing factors to the placebo response, published in English between January 1990 and December 2017. Results were filtered to only show meta-analyses, reviews, and systematic reviews. Relevant abstracts were hand-searched, full articles obtained, and information from these utilized to synthesize the present systematic review. Reference lists of articles were also examined to identify further relevant studies not identified by the keyword searches. Meta-analyses that aimed to evaluate the association of study features with placebo response or the differential response to antidepressants and placebo were included in the current review, provided they were based on 'statistical aggregation' of (patient-level) data or (study-level) results from placebo-`controlled RCTs in depression. To be included, underlying RCTs were required to have enrolled patients with depressive symptoms, fulfilling further diagnostic criteria of MDD according to the Diagnostic and Statistical Manual of Mental Disorders (DSM, version III, III-R, IV, or IV-TR) or Research Diagnostic Criteria (RDC), and assessed with commonly accepted primary outcome variables such as the Hamilton Rating Scale for Depression (HAMD, 17 or 21-item version), Montgomery & Åsberg Depression Rating Scale (MADRS), and/or Clinical Global Impression scale (CGI, severity and/or improvement).[16-23]

In order to evaluate the predictive strength of study outcomes, we assessed whether the meta-analyses (*1*) were based on a representative sample of RCTs, (*2*) focused on *illness severity* (improvement, or mean change in symptoms), or *curative effect* (percentage of participants fulfilling criteria for 'response' or 'remission') on placebo, rather than *trial outcome* (i.e., drug superiority over placebo, expressed in percentage of positive trials, or standardized mean difference between treatments) as endpoint, (*3*) applied formal and appropriate statistical testing, and (*4*) adhered to basic quality principles for meta-analysis. These four assessments are further explained below.

Ad (1). As far as could be verified, most of the meta-analyses were based on a sample of RCTs from two large and partially overlapping data sets, that were included in two studies by

Khan et al. (2010) and Furukawa et al. (2016). The two reference papers analyzed a total of 314 RCTs, testing the antidepressant qualities of 49 different drug formulations against placebo between the years 1978 and 2015.[15,24] For each meta-analysis, the amount of underlying RCTs already listed in the two reference papers was used to calculate the Jaccard index (*T*) as a measure of overlap or representativeness, using the formula:

$$T = N_c / (N_a + N_b - N_c)$$

whereby N_a is the total number of underlying RCTs included in the meta-analysis, N_b is the total number of RCTs listed in the two reference papers (N_b=314), and N_c is the number of RCTs in the meta-analysis that were also included in the two reference papers. When authors of a paper did not provide further details on RCTs underlying their meta-analysis, the maximum Jaccard index was calculated, assuming that all of the underlying RCTs already were included in the list of reference trials. In addition to the Jaccard index, the total number of trial participants exposed to placebo or active drug were extracted and tabulated, as well as the period in which underlying RCTs were completed or reported (whichever was mentioned).

Ad (2). For those meta-analyses in which a positive trial outcome or effect size (the difference between active drug and placebo) was used as an endpoint (rather than cure, or illness severity changes on placebo), results were considered to not provide direct evidence for an effect on placebo response.

Ad (3). Associations between explored variables and placebo response were only considered to be positive or negative when statistically significant under the reported testing conditions.

Ad (4). To further assess the scientific rigor of the meta-analyses, similar criteria were applied as on the basis of which the Overview Quality Assessment Questionnaire (OQAQ) was earlier validated as an index of the quality of review articles.[25] More specifically, this concerned the following criteria: (a) unpublished studies are included or searched for (comprehensive search); (b) search terms are clearly specified (selection bias was avoided); (c) descriptive data are presented for each study, such as design, subject characteristics, intervention, outcome, year(s) of study conduct or publication of results (validity of studies was assessed); (d) level of heterogeneity is explored and controlled for (studies were combined appropriately). As a fifth criterion, we added: (e) data were derived at patient-level, assuming a higher level of accuracy of findings in comparison with study-level data extraction, as explained in the introduction.

Procedures

A systematic extraction form was used to collect core details from all meta-analyses, including evidence for patient-, and trial variables that were somehow found to be associated with improvement on placebo or to have no apparent impact on placebo response. For the sake of convenience, predictors showing positive associations were color-coded in green, whereas those with negative associations were color-coded in red, and those without any significant relationship in blue. All predictors explored in the meta-analyses were listed using pre-defined, categorical terms for investigated aspects of demographics, illness severity at baseline, illness duration, diagnostic subtype, study enrolment criteria, dosing strategy, visit schedule, primary outcome variable, RCT selection process, patient recruitment, and analysis methods. Through the elimination of differences in variable names, the total number of explored predictors was thus kept to a minimum, and cross-comparison of results between meta-analyses was facilitated.

As a second step, all 314 RCTs from the two reference papers were listed and amended with design details, as provided by Undurraga and Baldessarini (2012) for many of these, in an excel spreadsheet. Based on the information contained in each meta-analysis paper, the total number of RCTs included, in combination with the number of RCTs matching with RCTs in the reference list were used for automated calculation of the Jaccard index.

As a third step, results were tabulated and compared for all meta-analyses, counting the number of papers suggesting a positive-, negative-, or zero-effect association of a specific factor with placebo response. The overall direction of the association was determined on the basis of following terms (in decreasing order of importance): positive or negative associations on the basis of a patient-level analysis, positive or negative associations on the basis of large to very large samples, and the number of counts for any (or absence of) relationship.

As a fourth step, the level of evidence for each factor's overall positive, negative, or zero-effect (i.e., absence of) association with placebo response was determined. We first assessed whether there was support for the existence of any association that was (*1*) *substantial* (i.e., findings suggesting a positive, zero, or negative effect on placebo response, based on an analysis at patient-level *or* involving a representative sample); (*2*) *convincing* (i.e., findings suggesting a positive, zero, or negative effect on placebo response, based on at least one patient-level analysis *and* at least one analysis involving a large, representative sample); (*3*) *consistent* (i.e., findings in one direction only, or only findings suggesting a zero association

with placebo response); (*4*) *contradictory* (i.e., findings in two directions, suggesting both a positive and negative association with placebo response). In order to reduce bias, we subsequently used custom decision criteria to draw conclusions about the level of evidence, as suggested by the availability of substantial, convincing, consistent, and/or contradictory evidence. The decision criteria are available as Supplement I to this paper.

Finally, the collective support for factors to be associated with placebo response was gathered and compared for patient-level meta-analyses only, which used change in illness severity, response, or remission on placebo as dependent variable (from here called 'patient-level response evaluation'). For that purpose, the total sum of the Jaccard indices of all studies in support of a particular association was taken as a measure for the level of evidence.

For the first two steps, the extraction of information from all meta-analysis papers was done by the first author (JS), and from a random selection of 25% of the papers by the second author (SK). Factor retrieval, color-coding of detected associations, and calculations of the Jaccard index from both authors were entirely consistent, and without discrepancies. At that point it was decided, that duplicate extraction of information from all the remaining meta-analysis papers by a second investigator was not necessary due to the high inter-investigator reliability. For steps three and four, both investigators (JS and SK) assessed the overall direction of all identified factor associations with placebo response, as well as the level of evidence present for each association. Again, there were no differences between the investigators' judgments, so that no further measures were deemed necessary to control for potential bias in drawing up conclusions about the overall strength of evidence for particular factors influencing placebo response.

Since none of the pooled analyses that we included could be regarded as a meta-analysis in its strict sense, and analytical approaches varied considerably, the current work is descriptive only, and without formal statistical analysis.

RESULTS

The search in PubMed yielded 1213 hits, from which titles were reviewed for relevant papers. Based on study titles alone, 38 studies were pre-identified. After reading the abstracts, ten studies were excluded from further consideration, because five of them summarized previous work without presenting new results from the further analysis, one presented results from a

single study and four failed to investigate factors possibly associated with placebo response. Full texts of the remaining 28 studies were obtained and further examined for their relevance to the topic under investigation. Ten studies were excluded from further consideration because of not fulfilling the selection criteria for pooled analysis and/or quality of underlying study results. The examination of the reference lists of the remaining 18 studies yielded an additional 40 studies, resulting in a total of 58 studies included in the current systematic review (Figure 2.1).[1,9-15,24,26-74] The vast majority of identified meta-analyses (75%) were published between 2005 and 2015. Details of all 58 studies are provided Table 2.1.

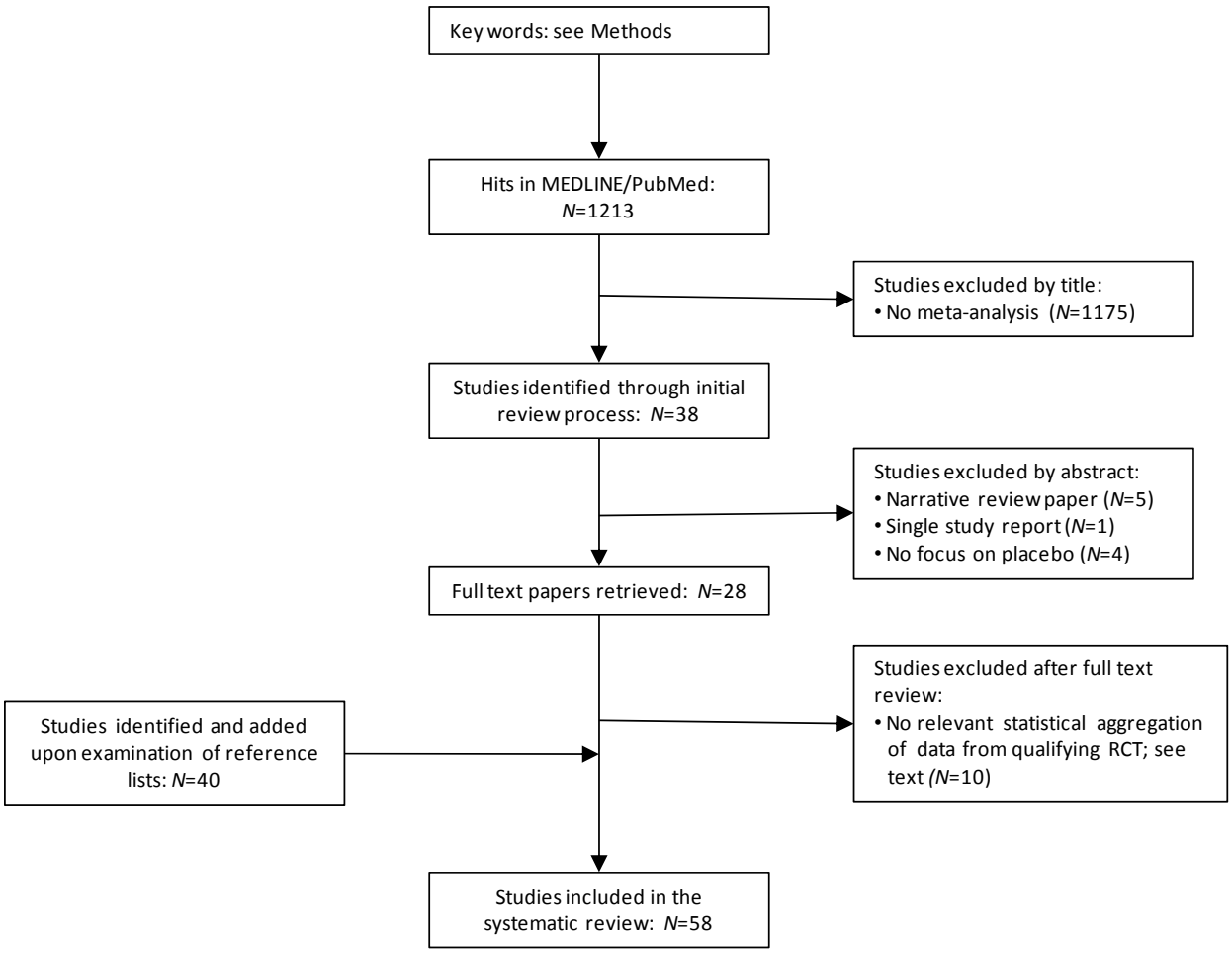


Figure 2.1

Study selection procedure

The effects of 40 independent variables on placebo response were explored. These variables can be subdivided into: (*1*) patient characteristics (demographics, illness severity,

illness duration, diagnostic subtype); (*2*) trial design features (enrolment criteria, dosing strategy, visit schedule, primary outcome variable); and (*3*) data sampling and handling procedures (trial selection, patient recruitment, analysis method).

Three different types of dependent variables were used in the meta-analyses, all with the aim to predict a potential impact of specific patient-, trial-, sampling-, or analysis features on placebo response. Sixteen meta-analyses focused on improvement on placebo versus drug as outcome variable (i.e., change in illness severity), 25 studies focused on response or remission on placebo (i.e., a 'curative effect'), and 17 on drug superiority over placebo or standardized mean difference between treatments (i.e., trial outcome). Although the latter can only yield indirect evidence of a factor's effect on placebo response, and '(non-)response' or '(non-) remission' as a dichotomous variable is less informative and sensitive than other categories of variables, such as mean difference in change from baseline scores on placebo versus drug, none of the authors provided an adequate rationale for their chosen outcome variable. Also, the applied statistical analysis methods varied widely, and even though appropriately applied, were mostly selected without justification.

Almost 60% (N=34) of the analyses were done at study-level, involvng 12 to 252 RCTs. In addition, 24 analyses were done at patient-level, most of which were based on a sample of fewer than 10 RCTs. None of the meta-analyses fulfilled all five quality criteria, and most of them involved a limited sample of non-unique RCTs, resulting in very low to relatively low Jaccard index values (between 0.00 and 0.05 in the patient-level analyses).

Suggested associations between explored parameters and placebo response from all meta-analyses are listed in Table 2.2. Convincing findings, based on at least one analysis at patient-level (in bold) *and* one involving a large, representative sample (indicated by **) were consistently recorded for a variety of factors, providing strong to very strong evidence for *absence* of an effect of age, sex, placebo run-in, region (% US vs. non-US sites), and trial completion year on the magnitude of placebo response. Substantial findings, based on patient-level analyses *or* a large representative sample, were obtained for the vast majority of factors but appeared only consistent (i.e., unidirectional) in a small, scarcely investigated, minority, providing moderate evidence for a positive (enhancing) effect of structured vs. non-structured interviews and number of participating sites per study, and *absence* of an effect of body-mass index on the magnitude of placebo response.

The level of evidence for a positive or negative association with placebo response was absent, or very weak to weak for all other factors explored, either because of contradictory or merely isolated findings (Table 2.2).

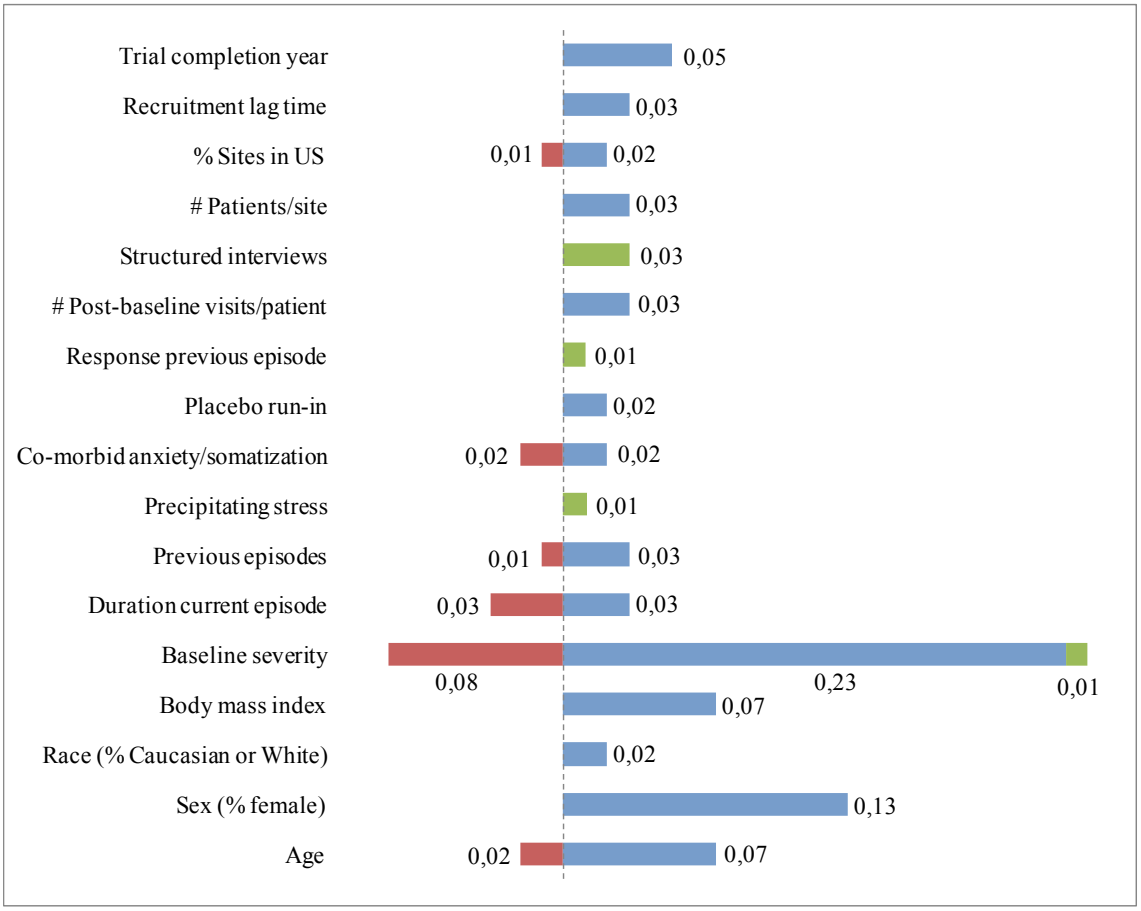


Figure 2.2

Cumulative Jaccard index counts of patient-level meta-analyses, suggesting a positive (green), negative (red) or zero (blue) association between patient-, trial-, or sampling-related factors on the one hand, and placebo response on the other

A comparable overall pattern of lack of associations with placebo response was also found in the patient-level response evaluation (Figure 2.2). Interpretation of the data in this subset of studies is hampered by the relatively few RCTs they were based upon, as reflected by the low, cumulative Jaccard indices. An early meta-analysis of results from 241 patients in three RCTs, led Brown et al. (1992) to conclude that precipitating stress and response in previous episodes are positively associated with placebo response.[26] Khan et al. (2014a, 2014b) claimed to have found evidence of a positive association between placebo response and use of

structured clinical interviews in their analyses of results from a total of 221 patients treated in a single center.[71,72] In all cases where support was found for a variable to be negatively associated with placebo response, there was at least equal support found for the same variable to be not at all associated with placebo response.

DISCUSSION

The present systematic review of previous meta-analyses for the first time brings together all collective evidence of factors putatively influencing placebo response in depression trials. Although there is potential overlap in underlying RCTs included, the meta-analyses can be regarded as more or less independent evaluations because of differences in chosen predictor and outcome variables, study populations, and heterogeneity in RCT design. Our results indicate that the level of evidence for a positive or negative association with placebo response is very weak to weak at best for all evaluated factors.

Although our findings suggest that there is moderate evidence for a positive effect on placebo response of structured (vs. non-structured) interview techniques and total number of sites/trial, these findings are based on two very small patient-level meta-analyses of single-center data focusing on illness severity (structured interviews), and three study-level analyses focusing on cure as endpoint (sites/study). One should be cautious to conclude that the effect of structured interview techniques on placebo response, as determined by Khan et al. in a selected sample of 221 RCT participants from a single center, is representative of what can be expected in the general population of depressed patients, as long as the observation has not been replicated in a multi-regional trial.[71,72] Similarly, one should be cautious to conclude that a higher number of sites per RCT, as found by Bridge, Undurraga, Furukawa, and coworkers, would directly increase the likelihood of a response to placebo in an individual participant, unless this observation is replicated in patient-level analyses.[15,50,64] Furthermore, it remains to be clarified whether an increased placebo response in RCTs with a relatively high number of sites is not an artifact associated with other factors potentially influencing placebo response, such as competitive recruitment and other efforts for more rapid patient enrolment, whereby sponsors and/or investigators may trivialize the potential risks, and stress the potential benefits of double-blind treatment to participating patients.

A positive association between low illness severity at baseline and placebo response was reported in two out of three narrative meta-reviews of a series of meta-analyses (see Introduction above).[6-8] However, caution is needed when comparing and interpreting results from the various meta-analyses, as we shall briefly illustrate for the putative role of illness severity below. Factors' associations with placebo response may be strongly influenced by (1) choice of the study outcome variable, (2) sampling bias in the underlying RCT, (3) whether study-level or patient-level data are extracted, (4) whether heterogeneity in RCT design and duration, as well as patient characteristics, are examined and controlled for, and (5) methods used to analyze the data. In what follows, we will elaborate on each of these aspects.

First, as pointed out by Papakostas and coworkers, the degree or probability of improvement as a function of any patient characteristic or trial feature may vary according to whether improvement is defined as a continuous measure (favoring patients with more severe symptoms, as they may demonstrate a numerically greater reduction in scores), or dichotomous one (favoring patients with milder symptoms, when they require a smaller degree of symptom reduction until being considered improved or no longer ill).[6] On the other hand, using the relative efficacy of antidepressants versus placebo as outcome measure (e.g., effect size) can only provide indirect evidence of a factor being associated with placebo response, since observations may be confounded by an enhanced or impaired response to active treatment.

Second, sampling bias may play a role, for example, whether the influence of baseline severity on placebo response is explored in severely or mildly ill patients. Study-level analyses of 52 RCTs by Khan et al. (2002, 2004) and 35 RCTs by Kirsch et al. (2008) suggested reduced placebo response with increased illness severity (i.e., a negative association). However, the RCTs they had included in their analysis almost exclusively comprised samples of patients with a mean baseline $HAMD_{17}$ score $\geq$23, i.e., very severe depression.[31,47] Melander and coworkers (2008), in contrast, concluded on the basis of a study-level analysis of 56 RCTs submitted to the European regulatory authorities comprising samples of patients with a relatively low enrolment threshold ($HAMD_{17}$ score $\geq$15) that there was no statistical evidence of a relation between average baseline $HAMD_{17}$ scores and percentage 'responders' on active treatment or placebo. Results from Fournier et al. (2010) may partly explain the discrepancy in findings, suggesting that placebo response (and even more so the response to the active agent) increases with increasing baseline HAMD score in

patients with mild to severe depression, but that the effect of placebo wanes in patients with very severe depression. Thus, some observations could reflect nothing more than just a regression to the mean, whereby the response to placebo becomes (relatively more strongly) attenuated by treatment resistance in patients with very severe depression.[53]

Third, as long as meta-regression models rely on study averages, conclusions about the potential influence of patient characteristics on outcome remain crude and potentially influenced by variability in trial design and data acquisition methods. In a study-level analysis by Khan and co-workers (2004) of 52 RCTs submitted to the FDA, an increased baseline illness severity was significantly associated with an increased difference in HAMD change scores between active drug and placebo, suggesting a *decreased* response to placebo.[9] Nevertheless, the same authors were unable to confirm this in a patient-level analysis of data from 15 RCTs at their site one year later.[38] Several years later, using more or less the same study-level data in a hierarchic multiple regression model, Khan and co-workers (2007) were unable to demonstrate a statistically significant association between baseline illness severity on the one hand, and difference in HAMD change scores between active drug and placebo on the other.[41] In contrast, when using study-level data from 130 RCTs that were published between 1981 and 2008 using hierarchic multiple regression models, Khan and co-workers (2010) found a statistically significant association between increased baseline illness severity and *increased* improvement on placebo.[24] However, as earlier, Khan and co-workers (2011) were unable to demonstrate the presence of a statistically significant association between baseline illness severity and improvement on placebo in a multiple linear regression model using patient data from 15 RCTs conducted at their site between 1995 and 2004.[54]

Fourth, in a conventional pooled-analysis, participant data from various studies with a common aim and design are shared to produce combined datasets that are subsequently analyzed as a single study. In a conventional meta-analysis, study-level effect estimates are averaged and aggregated to approximate the overall effect - for all studies combined - in the investigated treatment setting. In the 58 papers included in the current systematic review, neither of these two approaches was used. Instead, analogous to the conventional technique of statistical inference, data from various RCTs were integrated and properties deduced from the underlying distribution by analysis of the data in accordance with basic assumptions about heterogeneity, sampling, distribution, and covariance of relevant parameters in the chosen statistical model. Unfortunately, because none of the studies were designed to investigate the

association of study features with placebo response, the power of the statistical techniques applied in the reviewed publications is rather low and the risk for erroneous conclusions high. This may explain why there were no associations between illness severity and placebo response in 18 meta-analyses, whereas a negative association was suggested eight times, and a positive relationship four times.

Fifth, conclusions from Kirsch and coworkers (2008) about the presence of an association between baseline illness severity and placebo response relied on randomized cohorts rather than randomized individuals. A re-analysis of these data, fitting random effects models in both Bayesian and frequentist statistical frameworks (using raw mean difference and standardized mean difference scales) implied, however, that there is no significant role of baseline illness severity in treatment outcome.[69] This conclusion is further supported by the lack of evidence for any association between baseline illness severity and placebo response from a recent patient-level meta-analysis of 34 RCTs.[74]

As far as we are aware, this is the first exhaustive 'meta-review' of published meta-analyses investigating factors contributing to placebo response in depression RCTs, whereby findings in support of no association were included rather than ignored, and the level of evidence for any contributing effect was systematically explored for each factor on the basis of pre-defined criteria. The main strength of the present study, i.e., the use of pre-defined criteria for the systematic weighing of evidence in support of an association between explored features and placebo response, may at the same time be regarded as a serious weakness. The criteria have been set to increase objectivity when comparing the results of a wide variety of analyses. However, these criteria have not yet been validated. When evidence is labeled as 'convincing' or 'substantial,' this does not take into account the accuracy or reliability of the analysis and observations. It may thus be that the evidence is considered meaningful, although the underlying research is of relatively poor quality. Our separate patient-level response evaluation aimed to remove noise coming from studies that used (1) aggregate study data or (2) trial outcome as dependent variable. Results were more or less comparable with those from the overall evaluation, although based on studies which used far less underlying RCTs. The cumulative Jaccard indices in this approach appeared to be very low, and may in some cases even be overestimates of 'evidence' being present, since the individual values are strictly spoken not additive, especially not when studies are partly based on the same RCTs or apply different statistical models. Although inclusion of the Jaccard index as a measure of the

strength of evidence may introduce bias against older studies (since fewer RCTs were available at the time of the analysis), we felt that only the results of studies based on a representative sample of RCTs may have likely implications for placebo-controlled, clinical trials. The 'representativeness' of an underlying sample of RCTs was defined by a Jaccard index value of twice the median (i.e., more than 0.256). This truncation value was rather arbitrarily chosen. However, when set to 'above median' (i.e., more than 0.129), the strength of evidence for all predictors remains the same except for concomitant medication and published vs. unpublished RCTs (both changing from 'very weak' to 'weak' evidence). The Jaccard index itself, of which the values vary by definition from 0.00 (no overlap) to 1.00 (full overlap) has to be interpreted with caution because sample sizes of underlying RCTs may vary substantially, and results of a meta-analysis of a selection of large RCT can be more informative than when based on a selection of small RCT, while the Jaccard index may be the same. Furthermore, a high volume of data does not necessarily predict high-quality output or absolute reliability of results. Several problems connected with meta-analysis remain unaddressed in this review. Inappropriate statistical techniques may sometimes have been used. For instance, a fixed-effects model is appropriate for study-level meta-analyses when all included RCTs are identical, and all observed variation is caused by chance or within-study sampling error. However, whenever there is an interest to generalize the results, and not all RCTs are of identical design and conduct, a random-effects model would be more appropriate.[75] As long as we do not know which patient characteristics or trial aspects can influence placebo response, the choice of either method is somewhat arbitrary. The approach of using pooled analyses of RCTs for evaluating factors other than (drug or placebo) treatment associated with response, ignores the fact that the principle of randomization is aimed to address the clinical question *whether* a treatment is efficacious and sufficiently well tolerated, and *not* in which patients or under which conditions. The validity of any association seemingly present between covariates and placebo response will, therefore, need to be replicated in large, well-designed, prospective, trials.

In conclusion, our systematic review demonstrates that meta-analyses published during the past 25 years did not answer questions about potential ways to avoid placebo response and trial failure in depression RCT. Despite frequently acclaimed associations between placebo response and specific patient or trial features, such as the probability of receiving placebo, increased visit frequency, the number of sites/trial, baseline illness severity, dosing regimen

and year of trial conduct, the level of evidence for these associations remains very weak to weak at best. There is stronger evidence for a plethora of factors not to be associated with placebo response. As long as there is no broad access to the individual patient data from RCTs, we recommend that future research includes the design of prospective RCTs evaluating placebo response in subpopulations treated under specific conditions. The medical-ethical challenges associated with such an approach may be prohibitive, however, and trial sponsors will have to continue dealing with high placebo response rates in a rather crude way, e.g., by increasing sample sizes, until it is sufficiently clear which patients, investigators, or study procedures should be avoided in order to minimize the likelihood of placebo response in depression trials.

REFERENCES

1. Montgomery SA, Kasper S. Severe depression and antidepressants: focus on a pooled analysis of placebo-controlled studies on agomelatine. *International Clinical Psychopharmacology* 2007;22(5):283-291.
2. Sysko R, Walsh BT. A systematic review of placebo response in studies of bipolar mania. *Journal of Clinical Psychiatry* 2007;68(8):1-478.
3. Kemp AS, Schooler NR, Kalali AH, Alphs L, Anand R, Awad G, Davidson M, Dubé S, Ereshefsky L, Gharabawi G. What is causing the reduced drug-placebo difference in recent schizophrenia clinical trials and what can be done about it? *Schizophrenia Bulletin* 2010;36(3):504-509.
4. Ackerman DL, Greenland S. Multivariate meta-analysis of controlled drug studies for obsessive-compulsive disorder. *Journal of Clinical Psychopharmacology* 2002;22(3):309-317.
5. Welten C, Koeter M, Wohlfarth T, Storosum J, van den Brink W, Gispen-de Wied C, Leufkens H, Denys D. Placebo response in antipsychotic trials of patients with acute mania: Results of an individual patient data meta-analysis. *European Neuropsychopharmacology* 2015;25(7):1018-1026.
6. Papakostas GI, Ostergaard SD, Iovieno N. The nature of placebo response in clinical studies of major depressive disorder. *Journal of Clinical Psychiatry* 2015;76(4):456-466.
7. Rutherford BR, Roose SP. A model of placebo response in antidepressant clinical trials. *American Journal of Psychiatry* 2013;170(7):723-733.
8. Weimer K, Colloca L, Enck P. Placebo effects in psychiatry: mediators and moderators. *The Lancet Psychiatry* 2015;2(3):246-257.
9. Khan A, Kolts RL, Thase ME, Krishnan KRR, Brown W. Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. *American Journal of Psychiatry* 2004;161(11):2045-2049.
10. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *European Neuropsychopharmacology* 2009;19(1):34-40.
11. Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lanctot KL. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses. *Journal of Clinical Psychiatry* 2010;71(3):270-279.
12. Iovieno N, Papakostas GI. Correlation between different levels of placebo response rate and clinical trial outcome in major depressive disorder: a meta-analysis. *Journal of Clinical Psychiatry* 2012;73(10):1300-1306.
13. Dunlop BW, Thase ME, Wun CC, Fayyad R, Guico-Pabia CJ, Musgnung J, Ninan PT. A meta-analysis of factors impacting detection of antidepressant efficacy in clinical trials: the importance of academic sites. *Neuropsychopharmacology* 2012;37(13):2830-2836.
14. Mancini M, Wade AG, Perugi G, Lenox-Smith A, Schacht A. Impact of patient selection and study characteristics on signal detection in placebo-controlled trials with antidepressants. *Journal of Psychiatric Research* 2014;51:21-29.
15. Furukawa TA, Cipriani A, Atkinson LZ, Leucht S, Ogawa Y, Takeshima N, Hayasaka Y, Chaimani A, Salanti G. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *Lancet Psychiatry* 2016;3(11):1059-1066.
16. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders DSM-III (3rd.ed.).* Washington, DC1980.
17. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders DSM-III-R (3rd ed., rev.).* Washington, DC1987.
18. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders DSM-IV (4th ed.).* Washington, DC1994.
19. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision).* Washington, DC: American Psychiatric Association; 2000.
20. Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* 1978;35(6):773-782.
21. Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry* 1960;23(1):56-62.
22. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* 1979;134(4):382-389.
23. Guy W. Clinical global impression scale. *The ECDEU Assessment Manual for Psychopharmacology-Revised Volume DHEW Publ No ADM* 1976;76(338):218-222.

24. Khan A, Bhat A, Kolts R, Thase ME, Brown W. Why has the antidepressant-placebo difference in antidepressant clinical trials diminished over the past three decades? *CNS Neuroscience & Therapeutics* 2010;16(4):217-226.
25. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology* 1991;44(11):1271-1278.
26. Brown WA, Johnson MF, Chen M-G. Clinical features of depressed patients who do and do not improve with placebo. *Psychiatry Research* 1992;41(3):203-214.
27. Trivedi MH, Rush J. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology* 1994;11(1):33-43.
28. Greenberg RP, Fisher S, Riter JA. Placebo washout is not a meaningful part of antidepressant drug trials. *Perceptual and Motor Skills* 1995(81):688-690.
29. Faries D, Herrera J, Rayamajhi J, DeBrota D, Demitrack M, Potter WZ. The responsiveness of the Hamilton depression rating scale. *Journal of Psychiatric Research* 2000;34(1):3-10.
30. Entsuah R, Shaffer M, Zhang J. A critical examination of the sensitivity of unidimensional subscales derived from the Hamilton Depression Rating Scale to antidepressant drug effects. *Journal of Psychiatric Research* 2002;36(6):437-448.
31. Khan A, Leventhal RM, Khan SR, Brown WA. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *Journal of Clinical Psychopharmacology* 2002;22(1):40-45.
32. Quitkin FM, Stewart JW, McGrath PJ, Taylor BP, Tisminetzky MS, Petkova E, Chen Y, Ma G, Klein DF. Are there differences between women's and men's antidepressant responses? *American Journal of Psychiatry* 2002;159(11):1848-1854.
33. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* 2002;287(14):1840-1847.
34. Khan A, Khan SR, Walens G, Kolts R, Giller EL. Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the food and drug administration summary basis of approval reports. *Neuropsychopharmacology* 2003;28(3):552-557.
35. Stolk P, ten Berg MJ, Hemels ME, Einarson TR. Meta-analysis of placebo rates in major depressive disorder trials. *Annals of Pharmacotherapy* 2003;37(12):1891-1899.
36. Lee S, Walker JR, Jakul L, Sexton K. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? A meta-analytic evaluation. *Depression and Anxiety* 2004;19(1):10-19.
37. Evans KR, Sills T, Wunderlich GR, McDonald HP. Worsening of depressive symptoms prior to randomization in clinical trials: a possible screen for placebo responders? *Journal of Psychiatric Research* 2004;38(4):437-444.
38. Khan A, Brodhead AE, Kolts RL, Brown WA. Severity of depressive symptoms and response to antidepressants and placebo in antidepressant trials. *Journal of Psychiatric Research* 2005;39(2):145-150.
39. Stein DJ, Baldwin DS, Dolberg OT, Despiegel N, Bandelow B. Which Factors Predict Placebo Response in Anxiety Disorders and Major Depression? *Journal of Clinical Psychiatry* 2006;67(11):1741-1746.
40. Lam RW, Andersen HF. The influence of baseline severity on efficacy of escitalopram and citalopram in the treatment of major depressive disorder: an extended analysis. *Pharmacopsychiatry* 2006;39(5):180-184.
41. Khan A, Schwartz K, Kolts RL, Ridgway D, Lineberry C. Relationship between depression severity entry criteria and antidepressant clinical trial outcomes. *Biological Psychiatry* 2007;62(1):65-71.
42. Posternak MA, Zimmerman M. Therapeutic effect of follow-up assessments on antidepressant and placebo response rates in antidepressant efficacy trials Meta-analysis. *British Journal of Psychiatry* 2007;190(4):287-292.
43. Mallinckrodt CH, Meyers AL, Prakash A, Faries DE, Detke MJ. Simple options for improving signal detection in antidepressant clinical trials. *Psychopharmacology Bulletin* 2007;40(2):101.
44. Entsuah R, Vinall P. Potential predictors of placebo response: Lessons from a large database. *Drug Information Journal* 2007;41(3):315-330.
45. Thase ME, Pritchett YL, Ossanna MJ, Swindle RW, Xu J, Detke MJ. Efficacy of duloxetine and selective serotonin reuptake inhibitors: comparisons as assessed by remission rates in patients with major depressive disorder. *Journal of Clinical Psychopharmacology* 2007;27(6):672-676.

46. Melander H, Salmonson T, Abadie E, van Zwieten-Boot B. A regulatory Apologia--a review of placebo-controlled studies in regulatory submissions of new-generation antidepressants. *European Neuropsychopharmacology* 2008;18(9):623-627.
47. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 2008;5(2):0260-0268.
48. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine* 2008;358(3):252-260.
49. Brunoni AR, Lopes M, Kaptchuk TJ, Fregni F. Placebo response of non-pharmacological and pharmacological trials in major depression: a systematic review and meta-analysis. *PLoS One* 2009;4(3):e4824.
50. Bridge JA, Birmaher B, Iyengar S, Barbe RP, Brent DA. Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. *American Journal of Psychiatry* 2009;166(1):42-49.
51. Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Hofmann SG. Meta-analysis of the placebo response in antidepressant trials. *Journal of Affective Disorders* 2009;118(1-3):1-8.
52. Hunter AM, Cook IA, Leuchter AF. Impact of antidepressant treatment history on clinical outcomes in placebo and medication treatment of major depression. *Journal of Clinical Psychopharmacology* 2010;30(6):748-751.
53. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, Fawcett J. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;303(1):47-53.
54. Khan A, Bhat A, Faucett J, Kolts R, Brown WA. Antidepressant-placebo differences in 16 clinical trials over 10 years at a single site: role of baseline severity. *Psychopharmacology* 2011;214(4):961-965.
55. Tedeschini E, Levkovitz Y, Iovieno N, Ameral VE, Nelson JC, Papakostas GI. Efficacy of antidepressants for late-life depression: a meta-analysis and meta-regression of placebo-controlled randomized trials. *Journal of Clinical Psychiatry* 2011;72(12):1660-1668.
56. Khin NA, Chen Y-F, Yang Y, Yang P, Laughren TP. Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US Food and Drug Administration in support of new drug applications. *Journal of Clinical Psychiatry* 2011;72(4):1,478-472.
57. Iovieno N, Tedeschini E, Ameral VE, Rigatelli M, Papakostas GI. Antidepressants for major depressive disorder in patients with a co-morbid axis-III disorder: a meta-analysis of patient characteristics and placebo response rates in randomized controlled trials. *International Clinical Psychopharmacology* 2011;26(2):69-74.
58. Klemp M, Tvete IF, Gåsemyr J, Natvig B, Aursnes I. Meta-regression analysis of paroxetine clinical trial data: does reporting scale matter? *Journal of Clinical Psychopharmacology* 2011;31(2):201-206.
59. Rutherford BR, Sneed JR, Tandler JM, Rindskopf D, Peterson BS, Roose SP. Deconstructing pediatric depression trials: an analysis of the effects of expectancy and therapeutic contact. *Journal of the American Academy of Child & Adolescent Psychiatry* 2011;50(8):782-795.
60. Fountoulakis KN, Möller H-J. Efficacy of antidepressants: a re-analysis and re-interpretation of the Kirsch data. *International Journal of Neuropsychopharmacology* 2011;14(3):405-412.
61. Iovieno N, Tedeschini E, Levkovitz Y, Ameral VE, Papakostas GI. Does the frequency of follow-up assessments affect clinical trial outcome? A meta-analysis and meta-regression of placebo-controlled randomized trials. *International Journal of Neuropsychopharmacology* 2012;15(3):289-296.
62. Papakostas GI, Fan H, Tedeschini E. Severe and anxious depression: combining definitions of clinical sub-types to identify patients differentially responsive to selective serotonin reuptake inhibitors. *European Neuropsychopharmacology* 2012;22(5):347-355.
63. Nelson JC, Zhang Q, Deberdt W, Marangell LB, Karamustafalioglu O, Lipkovich IA. Predictors of remission with placebo using an integrated study database from patients with major depressive disorder. *Current Medical Research and Opinion* 2012;28(3):325-334.
64. Undurraga J, Baldessarini RJ. Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. *Neuropsychopharmacology* 2012;37(4):851-864.
65. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Archives of General Psychiatry* 2012;69(6):572-579.
66. Nelson JC, Delucchi KL, Schneider LS. Moderators of outcome in late-life depression: a patient-level meta-analysis. *American Journal of Psychiatry* 2013;170(6):651-659.

67. Rutherford BR, Cooper TM, Persaud A, Brown PJ, Sneed JR, Roose SP. Less is more in antidepressant clinical trials: a meta-analysis of the effect of visit frequency on treatment response and dropout. *Journal of Clinical Psychiatry* 2013;74(7):703-715.
68. Dodd S, Berk M, Kelin K, Mancini M, Schacht A. Treatment response for acute depression is not associated with number of previous episodes: lack of evidence for a clinical staging model for major depressive disorder. *Journal of Affective Disorders* 2013;150(2):344-349.
69. Fountoulakis KN, Veroniki AA, Siamouli M, Moller H-J. No role for initial severity on the efficacy of antidepressants: results of a multi-meta-analysis. *Annals of General Psychiatry* 2013;12:1-10.
70. Rutherford BR, Tandler J, Brown PJ, Sneed JR, Roose SP. Clinic Visits in Late Life Depression Trials: Effects on Signal Detection and Therapeutic Outcome. *American Journal of Geriatric Psychiatry* 2014;22(12):1452-1461.
71. Khan A, Faucett J, Brown WA. Magnitude of change with antidepressants and placebo in antidepressant clinical trials using structured, taped and appraised rater interviews (SIGMA-RAPS) compared to trials using traditional semi-structured interviews. *Psychopharmacology* 2014;231(22):4301-4307.
72. Khan A, Faucett J, Brown WA. Magnitude of placebo response and response variance in antidepressant clinical trials using structured, taped and appraised rater interviews compared to traditional rating interviews. *Journal of Psychiatric Research* 2014;51:88-92.
73. Locher C, Kossowsky J, Gaab J, Kirsch I, Bain P, Krummenacher P. Moderation of antidepressant and placebo outcomes by baseline severity in late-life depression: A systematic review and meta-analysis. *Journal of Affective Disorders* 2015;181:50-60.
74. Rabinowitz J, Werbeloff N, Mandel FS, Menard F, Marangell L, Kapur S. Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. *The British Journal of Psychiatry* 2016;209(5):427-428.
75. Kelley GA, Kelley KS. Statistical models for meta-analysis: a brief tutorial. *World Journal of Methodology* 2012;2(4):27-32.

Table 2.1

Meta-analyses of placebo response in depression RCT, published between 1992 and 2017

PUBLICATION		TRIALS			SUBJECTS			PERIOD	ANALYSIS							
									Variables		Methods	Quality				
	1st Author, yr	Na	Nc	T	Placebo N	Drug N	Total N	Yr – Yr	Outcome	Predictor		UT	ST	DD	PL	HE
1	Brown, 1992[26]	3	N/A	≤0.01	241	N/A	241	N/A	Cure	Patient characteristics	ANOVA; χ^2				★	
2	Trivedi, 1994[27]	101	37	0.10	N/A	N/A	N/A	1959-1990	Cure	Enrollment criteria	Random-effects model					
3	Greenberg, 1995[28]	28	N/A	≤0.09	N/A	N/A	1,089	1983-1992	Illness severity	Enrollment criteria	Independent samples (Student's t-test)					
4	Faries, 2000[29]	12	N/A	≤0.04	N/A	N/A	2,899	N/A	Trial outcome	Primary outcome variable	ANCOVA				★	
5	Entsuah, 2002[30]	8	4	0.01	446	1,599	2,045	N/A	Trial outcome	Primary outcome variable	ANCOVA	★		★	★	★
6	Khan, 2002[31]	45	28	0.08	2,736	5,145	7,881	1985-1997	Illness severity	Patient characteristics	Univariate correlation	★	★			
7	Quitkin, 2002[32]	8	2	0.01	443	522	965	1983-2000	Time to Cure	Patient characteristics	Cox proportional hazards regression	★			★	
8	Walsh, 2002[33]	75	75	0.24	N/A	N/A	N/A	1981-2000	Cure	Patient characteristics; Enrolment criteria; Trial duration; Trial period	Multiple regression		★			
9	Khan, 2003[34]	51	35	0.10	3,335	4,676	9,313	1985-1998	Trial outcome	Dosing strategy	χ^2; ANCOVA	★	★	★		
10	Stolk, 2003[35]	24	22	0.06	1,786	2,673	4,459	1985-2000	Cure	Trial period	Random-effects model		★			
11	Khan, 2004[9]	52	35	0.11	N/A	N/A	N/A	1985-2000	Trial outcome	Patient characteristics; Dosing strategy	Univariate correlation	★	★			
12	Lee, 2004[36]	42	28	0.09	3,740	3,047	6,787	1985-2000	Trial outcome	Enrollment criteria	Random-effects model		★	★		★
13	Evans, 2004[37]	4	N/A	≤0.01	283	274	557	N/A	Trial outcome	Enrollment criteria	Independent samples (Student's t-test)	★			★	

PUBLICATION		TRIALS			SUBJECTS			PERIOD	ANALYSIS			Quality				
					Placebo	Drug	Total		Variables		Methods					
	1st Author, yr	Na	Nc	T	N	N	N	Yr – Yr	Outcome	Predictor		UT	ST	DD	PL	HE
14	Khan, 2005[38]	15	N/A	≤0.05	143	186	329	1995-2003	Illness severity	Patient characteristics	Univariate correlation				★	
15	Stein, 2006[39]	5	4	0.01	738	1,254	1,992	2002-2004	Illness severity	Patient characteristics; Dosing strategy; Trial region	ANCOVA			★	★	
16	Lam, 2006[40]	3	2	0.01	398	923	1,321	2002-2004	Cure	Patient characteristics	ANCOVA			★	★	
17	Khan, 2007[41]	51	36	0.11	3,987	7,283	11,270	1985-2004	Illness severity	Patient characteristics; Enrollment criteria; Dosing strategy; Trial duration	Multiple regression		★	★		★
18	Posternak, 2007[42]	41	40	0.13	3,063	4,611	7,674	1981-2001	Illness severity	Visit schedule	No formal statistical testing					
19	Mallinckrodt, 2007[43]	11	5	0.02	1,053	2,357	3,410	N/A	Trial outcome	Enrollment criteria; Primary outcome variable; Analysis method	No formal statistical testing	★		★	★	
20	Entsuah, 2007[44]	18	8	0.02	1,651	2,850	4,501	1992-2006	Cure	Patient characteristics	Multiple logistic regression; χ^2	★		★	★	★
21	Thase, 2007[45]	6	6	0.03	516	1,317	1,833	2002-2006	Cure	Patient characteristics	χ^2			★	★	
22	Montgomery, 2007[1]	3	3	0.01	360	501	861	2002-2007	Illness severity	Patient characteristics	No formal statistical testing			★	★	★
23	Melander, 2008[46]	56	N/A	≤0.18	N/A	N/A	7,374	1984-2003	Cure	Patient characteristics	Linear regression		★			
24	Kirsch, 2008[47]	35	19	0.06	1,841	3,292	5,133	1987-1999	Trial outcome	Patient characteristics; Trial duration; Sample size	Meta-regression; Fixed-effects	★	★	★		
25	Turner, 2008[48]	74	49	0.14	5227	7337	12564	1987-2004	Trial outcome	Trial selection	Paired (signed-rank) analysis	★	★	★		
26	Brunoni, 2009[49]	41	31	0.10	2,394	2,682	5,076	2002-2008	Illness severity	Patient characteristics; Primary outcome variable	Multiple regression; Mixed-effects	★	★	★		★
27	Bridge, 2009[50]	12	0	0.04	N/A	N/A	2,862	1995-2004	Cure	Patient characteristics; Enrollment criteria; Sampling method; Trial region	Multiple regression		★	★		★

PUBLICATION		TRIALS			SUBJECTS			PERIOD	ANALYSIS							
									Variables		Methods	Quality				
	1st Author, yr	Na	Nc	T	Placebo N	Drug N	Total N	Yr – Yr	Outcome	Predictor		UT	ST	DD	PL	HE
28	Rief, 2009[51]	96	N/A	≤0.30	N/A	N/A	9,566	1980-2005	Illness severity	Patient characteristics; Trial duration; Visit schedule; Trial period	Multiple regression; Random-effects		★			
29	Papakostas, 2009[10]	182	N/A	≤0.57	13,107	23,278	36,385	1980-2007	Cure	Patient characteristics; Dosing strategy; Sampling method; Trial period	Multiple regression		★			★
30	Hunter, 2010[52]	3	0	0.01	35	37	72	N/A	Illness severity	Patient characteristics	ANCOVA; χ^2				★	★
31	Fournier, 2010[53]	6	1	0.00	284	434	718	1989-2009	Trial outcome	Patient characteristics	ANCOVA		★	★	★	
32	Sinyor, 2010[11]	35	30	0.09	N/A	N/A	N/A	1996-2007	Cure	Sampling method; Trial period	Poisson regression; Random-effects	★	★			
33	Khan, 2010[24]	125	63*	0.17	11,965	23,157	35,122	1981-2008	Illness severity	Patient characteristics; Dosing strategy; Sampling method; Trial period	Multiple regression	★	★			
34	Khan, 2011[54]	15	N/A	≤0.05	140	262	402	1995-2004	Illness severity	Patient characteristics; Trial period	Multiple regression			★	★	★
35	Tedeschini, 2011[55]	74	N/A	≤0.23	7,447	13,125	20,572	1980-2010	Cure	Patient characteristics	Meta-regression; Random-effects		★	★		★
36	Khin, 2011[56]	81	N/A	≤0.26	N/A	N/A	21,611	1983-2008	Cure / Trial outcome	Patient characteristics; Dosing strategy; Sampling method; Trial period	No formal statistical testing	★	★			
37	Iovieno, 2011[57]	197	N/A	≤0.63	17,319	29,581	46,900	1980-2009	Cure	Patient characteristics	ANCOVA		★			
38	Klemp, 2011[58]	26	15	0.05	2,123	2,958	5,081	1984-2006	Trial outcome	Patient characteristics; Sampling method; Trial period	Meta-regression; Random-effects	★	★	★		★
39	Rutherford, 2011[59]	18	0	0.06	1,509	2,373	3,882	1987-2009	Cure	Patient characteristics; Enrolment criteria; Sampling method	Logistic regression; Mixed-effects		★	★		★
40	Fountoulakis, 2011[60]	35	19	0.06	1,841	3,292	5,133	1985-1998	Trial outcome	Patient characteristics; Sampling method; Trial period	Meta-regression; Fixed-effects	★	★	★		
41	Iovieno, 2012[12]	176	N/A	≤0.56	16,871	2,437	29,437	1980-2011	Cure	Dosing strategy	Multiple regression; Random-effects		★			★

PUBLICATION		TRIALS			SUBJECTS			PERIOD	ANALYSIS			Quality				
					Placebo	Drug	Total		Variables		Methods					
	1st Author, yr	Na	Nc	T	N	N	N	Yr – Yr	Outcome	Predictor		UT	ST	DD	PL	HE
42	Iovieno, 2012[61]	146	N/A	≤0.46	13,949	24,911	38,860	1980-2010	Trial outcome	Sampling method	Multiple regression; Random-effects		★			★
43	Dunlop, 2012[13]	30	24	0.07	3,470	5,463	8,933	1993-2011	Cure / Trial outcome	Patient characteristics; Dosing strategy; Visit schedule; Trial region; Trial period	Multiple regression; Random-effects	★	★	★		★
44	Papakostas, 2012[62]	5	4	0.01	661	1,029	1,690	2002-2008	Trial outcome	Patient characteristics	ANCOVA			★	★	
45	Nelson, 2012[63]	8	5	0.02	1,328	2,132	3,460	2002-2007	Cure	Patient characteristics; Trial region	Multiple logistic regression & CART			★	★	
46	Undurraga, 2012[64]	107	94	0.29	17,059	9,925	27,127	1983-2010	Cure	Treatment strategy; Sampling method; Trial period	Linear regression		★	★		
47	Gibbons, 2012[65]	41	17	0.05	3,706	5,479	9.185	N/A	Illness severity	Patient characteristics	Linear regression; Random-effects	★		★	★	
48	Nelson, 2013[66]	7	7	0.02	994	1,494	2,488	2003-2010	Cure	Patient characteristics	Multiple logistic regression			★	★	
49	Rutherford, 2013-[67]	62	55	0.17	6,750	13,676	20,426	1985-2011	Cure	Visit schedule	Logistic regression; Mixed-effects		★	★		★
50	Dodd, 2013[68]	11	7	0.02	1,294	2,814	4,108	2000-2012	Cure	Patient characteristics	ANCOVA				★	
51	Fountoulakis, 2013[69]	35	19	0.06	1,841	3,292	5,133	1987-1999	Trial outcome	Patient characteristics	Meta-regression; Random-effects		★			★
52	Rutherford, 2014 [75]	19	11	0.03	1,854	3,555	5,409	1985-2010	Cure	Visit schedule	Multiple regression; Fixed-effects		★	★		★
53	Khan, 2014a[72]	7	N/A	≤0.02	162	0	162	2008-2012	Illness severity	Rating method	Levene's test			★	★	
54	Khan, 2014b[71]	4	N/A	≤0.01	59	89	148	2009-2011	Illness severity	Rating method	ANOVA			★	★	
55	Mancini, 2014[14]	14	10	0.03	1,573	3,088	4,661	1998-2007	Trial outcome	Patient characteristics; Dosing strategy; Visit schedule; Sampling method	Meta-regression; Fixed-effects	★	★	★	★	

	PUBLICATION	TRIALS			SUBJECTS			PERIOD	ANALYSIS			Quality				
									Variables		Methods					
	1st Author, yr	Na	Nc	T	Placebo N	Drug N	Total N	Yr – Yr	Outcome	Predictor		UT	ST	DD	PL	HE
56	Locher, 2015[73]	19	11	0.03	2,511	3,226	5,737	1980-2014	Trial outcome	Patient characteristics; Primary outcome variable	Meta-regression; Random effects		★	★		
57	Furukawa, 2016[15]	252	63**	0.13	26,323	N/A	N/A	1978-2015	Cure	Enrolment criteria; Dosing strategy; Sampling method; Trial period	Meta-regression; Random-effects	★				★
58	Rabinowitz, 2016[74]	34	15	0.05	3,258	7,323	10,581	1987-2007	Illness severity	Patient characteristics	Linear regression; Random- and fixed-effects	★			★	

* Listed by Furukawa et al. (2016) only;

** Listed by Kahn et al. (2010) only.

Yr = Year; N/A = Not available (no details provided by author); Na = Total number of RCTs included in the analysis; Nc = Number of RCTs out of the 315 (=Nb) RCTs listed by Khan et al (2010) and Furukawa et al (2016); Jaccard index: T = Nc / (Na+Nb-Nc); Yr = year; N = total number of patients on placebo, drug, (active control), or both; UT = Unpublished RCTs included in the analysis; ST = Search terms are clearly specified; DD = Descriptive data are presented for each underlying RCT (design, subject characteristics, intervention, outcome, year(s) of study conduct or publication of results); PL = Data extraction occurs at patient level; HE = Level of heterogeneity is explored and controlled for.

Table 2.2

Level of evidence for, and direction of association between placebo response and investigated factors

	Determined or inferred association with placebo response				
Predictor	**Negative**	**Zero**	**Positive**	**Net effect on Placebo response**	**Evidence**
PATIENT CHARACTERISTICS					
Demographics					
1. Age	**20**, 45	7, 8*, 26, 27, 28**, 29***, **34**, 35*, 38, 39, 43, **48**, 57***		Absent	Strong
2. Sex (% female)		**7**, 8*, **15**, **20**, 26, 27, 28**, **30**, **34**, 43, **45**, **48**, 57***	11	Absent	Very strong
3. Race (% Caucasian or White)		27, 43, **45**		Absent	Weak
4. Body mass index		**34**, **45**		Absent	Moderate
Illness severity					
5. Baseline HAMD or other scale	**1**, [11], **20**, **21**, [24], 29***, [36]**, **45**	6, 8*, **14**, **15**, **16**, 17, 23*, 26, 27, **30**, 38, 40, 41***, **47**, **48**, 51, **55**, 37***	**22**, **[31]**, 33***, [43]	Decrease	Very weak
6. Comorbidity				Absent	Weak
7. Concomitant medication		8*		Absent	Very weak
Illness duration					
8. Duration of illness	48	27		Decrease	Very weak
9. Duration current episode	**1**, **45**	8*, **15**, **20**		Decrease	Very weak
Diagnostic subtype					
10. Previous episodes	**30**	**1**, 27, 48, **50**		Decrease	Very weak
11. Precipitating stress			**1**	Increase	Weak

Predictor	Determined or inferred association with placebo response			Net effect on Placebo response	Evidence
	Negative	Zero	Positive		
12. Co-morbid axis-I condition:					
Anxiety/somatization	**45**	**1**, **44**		Decrease	Very weak
Sleep disturbance		**1**		Absent	Weak
Cognitive disturbance		**1**		Absent	Weak
Retardation		**1**		Absent	Weak
Melancholic features		**45**		Absent	Weak
Atypical features		**45**		Absent	Weak
13. Dysthymia	28**			Decrease	Weak
TRIAL DESIGN					
Enrolment criteria					
14. Placebo run-in		2, 3, 8*, 12, **19**, 27, 39, 57***		Absent	Very strong
15. Severity cut-off score		17, 57		Absent	Very weak
16. Recent change exclusion	**[13]**			Decrease	Very weak
17. Response previous episode			**1**	Increase	Weak
Dosing strategy					
18. Probability receiving placebo	[11], 29***, 32	41***, 43, 55, 57***		Decrease	Very weak
19. Dosing regimen (flexible vs. fixed)	[9], [11], [17]	15, 29***, 33***, 36**, 43	41***, 57**	Increase	Very weak
20. Study duration		8*, 17, 24, 27, 28*, 29***, 32, 38, **40**, 43	33***, 46**, 57**	Increase	Very weak

Predictor	Determined or inferred association with placebo response			Net effect on Placebo response	Evidence
	Negative	Zero	Positive		
Visit schedule					
21. Self-rating vs. observer-rating	28**			Decrease	Weak
22. Visits/Period	18, [42]***	39, 49*	52	Decrease	Very weak
23. Assessments/Visit		43		Absent	Very weak
24. Post-baseline visits/Patient		43, **55**		Absent	Weak
25. Structured vs. non-structured interview			**53**, **54**	Increase	Moderate
Primary outcome variable					
26. Uni- vs. multidimensional	**[4]**, **[5]**, **[19]**			Decrease	Very weak
27. MADRS vs. HAMD		26, 56		Absent	Very weak
28. $HAMD_{17}$ vs. $HAMD_{21}$		33***	[38]	Absent	Weak
DATA ACQUISITION & HANDLING					
Trial selection					
29. Published vs. unpublished	[25]*			Decrease	Very weak
30. Approved (vs. non-appr.) control		33***		Absent	Weak
Patient recruitment					
31. Patients/Trial		24, 27, 29***, 40, **41***, 57**	46**	Absent	Very weak
32. Patients/Site		27, **55**		Absent	Weak
33. Region (% US vs. non-US sites)	**15**	27, 36**, 43, **45**, 57***		Absent	Strong

Predictor	Determined or inferred association with placebo response: Negative	Zero	Positive	Net effect on Placebo response	Evidence
34. Patients/Arm		17	[36]**	Increase	Very weak
35. Sites/Study			27, 46**, 57***	Increase	Moderate
36. Academic vs. non-academic sites	[43]	43		Decrease	Very weak
37. Recruitment lag time		**55**		Absent	Weak
38. Trial period					
Trial start year		38, 43	36**	Increase	Weak
Trial completion year		**34**, 57***		Absent	Very strong
Trial publication year	28**	8, 32, 40, 41***	10, 29***, 33***, 46**	Increase	Very weak
Analysis method					
39. MMRM vs. LOCF	**[19]**			Decrease	Very weak
40. Completion rate		43		Absent	Very weak

Numbers refer to the sequential publication numbers (#) as listed in Table 2.1.

Publications of which the number is listed between [ ] show results in support of a possible effect of the predictor on placebo response in the direction indicated, either based on the predictor's association with the difference in mean change from baseline on placebo and active treatment (effect size), or trial success rate (statistical superiority of drug over placebo).

Figures in bold denote results from patient-level meta-analyses.

* Results based on the analyses of a representative sample of RCTs (T larger than the median value of 0.129)

** Results based on the analyses of a relatively large, representative sample of RCTs (T larger than twice the median value of 0.129)

*** Results based on the analyses of a very large, representative sample of RCTs (T larger than three times the median value of 0.129)

SUPPLEMENT 1

Decision criteria as used to establish level of evidence

Evidence	Support for a positive-, negative- or zero-effect association with placebo response
Very strong	Consistent, 'convincing' findings: (A) Patient-level analysis, *and* (B) Large representative sample (T > 0.25), *and* (C) No contradictory, 'substantial' findings, i.e. (A) or (B) does not apply for other effects
Strong	Consistent convincing findings, with one substantial contradictory finding: (A) *and* (B) apply, *and* (A) *or* (B) also applies for other effect-type
Moderate	Consistent substantial findings: (A) *or* (B) applies in ≥2 meta-analyses, *and* (A) *or* (B) does not apply for other effect-type
Weak	Skewed convincing, or isolated substantial findings: 1. Convincing positive- *or* negative-, *and* zero-effect finding(s), *or* 2. Isolated substantial finding
Very weak	Other skewed or above median findings: 1. Substantial or non-substantial positive- *or* negative-, *and* zero- effect finding(s), *or* 2. Above median convincing or substantial, positive- *or* negative-, *versus* opposite-effect findings, *or* 3. Above median non-substantial positive-, negative-, or zero-effect findings, *or* 4. Any of the above, based on drug superiority over placebo or trial success only

Chapter 3

Adjunctive medication for difficult to treat symptoms in schizophrenia

Evaluation of sustained, add-on treatment, while controlling for rater performance and indirect drug effects [1]

[1] This chapter was published as:

Schoemaker, J.H., Jansen, W.T., Schipper, J., & Szegedi, A. (2014).

The selective glycine uptake inhibitor Org 25935 as an adjunctive treatment to atypical antipsychotics in predominant persistent negative symptoms of schizophrenia: results from the GIANT trial.

Journal of Clinical Psychopharmacology 34(2), 190-198.

DOI: http://dx.doi.org/10.1097/JCP.0000000000000073

ABSTRACT

Background: Using a selective glycine uptake inhibitor as adjunctive to second-generation antipsychotic (SGA) was hypothesized to ameliorate negative and/or cognitive symptoms in subjects with schizophrenia.

Methods: Subjects with predominant persistent negative symptoms (previously stabilized ≥3 months on a SGA) were enrolled in a randomized, placebo-controlled trial to investigate adjunctive treatment with Org 25935, a selective inhibitor of the type-1 glycine transporter, over 12 weeks in a flexible dose design. Org 25935 was tested at 4-8 mg twice-daily and 12-16 mg twice-daily versus placebo. Primary efficacy outcome was the mean change from baseline in Scale for Assessment of Negative Symptoms ($SANS_{1-22}$) composite score. Secondary efficacy endpoints were the total and subscale scores on the Positive and Negative Syndrome Scale (PANSS), depressive symptoms (Calgary Depression Scale for Schizophrenia, CDSS), global functioning (Global Assessment of Functioning, GAF) and cognitive measures using a computerized battery (Central Nervous System Vital Signs). Responder rates were assessed post-hoc.

Results: A total of 215 subjects were randomized, of which 187 (87%) completed the trial. Both dose groups of Org 25935 did not differ significantly from placebo on SANS, PANSS (total or subscale scores), GAF, or the majority of tested cognitive domains. Org 25935 was generally well tolerated within the tested dose range, with no meaningful effects on EPS symptoms and some reports of reversible visual side effects.

Conclusion: Org 25935 did not differ significantly from placebo in reducing negative symptoms or improving cognitive functioning when administered as adjunctive treatment to SGA. In our study population Org 25935 appeared to be well tolerated in the tested dose ranges.

INTRODUCTION

In the treatment of schizophrenia, a substantial number of subjects present persistent negative symptoms and cognitive impairment, in many cases despite continued treatment with available antipsychotics.[1, 2] These symptom domains appear to account together for much of the long-term morbidity and poor functional outcome of subjects with schizophrenia.[3] There is a clear medical need to explore new options of effective treatments for these symptom clusters.

A growing body of evidence implies that alterations of glutamatergic neurotransmission may be of relevance in the neurobiology of schizophrenia, particularly in the domains of negative symptoms or cognitive impairment, in addition to dopaminergic dysregulations.[4] Multiple lines of evidence suggest a hypoactivity at the level of glutamatergic (especially N-methyl-D-aspartate [NMDA]) receptors in schizophrenia, which appear not to be addressed by currently available antipsychotics. A series of experimental drugs, aiming at a safe and tolerable increase of glutamatergic activity, have been investigated in schizophrenia, mainly as adjunctive therapy. A meta-analysis across studies with glycine, D-serine, D-cycloserine, and sarcosine suggests that NMDA-enhancing molecules may be effective in various symptom domains, with a combined effect size of up to 0.4 in depressive and negative symptoms, and 0.3 in positive and cognitive symptoms.[5]

The development of selective and potent inhibitors of the reuptake of glycine at the level of the glycine transporter type 1 (GlyT-1) has been suggested as a promising possibility to address currently unmet medical needs in the treatment of various symptom domains in schizophrenia, especially for negative symptoms or cognitive impairment.[6,7] Direct pharmacological activation of NMDA or other ionotropic glutamate receptors by an agonist is unlikely to be a useful approach because of the risk of overexcitation, which could cause neurotoxicity or seizures. Glycine binds to NMDA receptors and acts as a co-agonist for glutamatergic neurotransmission.[8] GlyT-1 inhibition is therefore regarded as a reasonable, non-excitotoxic approach for the enhancement of glutamatergic hypofunction associated with schizophrenia.

Org 25935 (cis - N - methyl- N - (6 - methoxy - 1 - phenyl -1,2,3,4- tetrahydronaphthalen-2 - ylmethyl) amino - methylcarboxylic acid hydrochloride) (Figure 3.1) is a highly selective inhibitor of GlyT-1 (IC_{50}=100 nM).[9] The molecular structure of the inhibitor contains a N-

methyl-glycine (sarcosine) moiety, suggesting glycine-competitive binding at the transporter binding site.[10] The compound has negligible effects on GlyT-2 and raises extracellular glycine levels in rat brain areas up to 2.3-fold after systemic administration.[11,12] Human glycine cerebro-spinal fluid (CSF) levels are increased in healthy volunteers up to about 2.5-fold after a single oral dose of 16 mg, whereas a smaller increase (about 1.5-fold) is found after a 4-mg dose, thereby providing evidence for relevant pharmacodynamic activity in the aforementioned dose range.[13]

In order to explore the therapeutic potential of Org 25935 in negative symptoms of schizophrenia, we conducted the Glycine uptake Inhibitor Add-on in Negative symptoms Trial (GIANT) in subjects stabilized on a second generation antipsychotic (SGA) with two adjunctive treatment regimens of Org 25935 and placebo. In addition to the effects on negative symptoms, we also aimed to explore the effects on positive symptoms, symptoms of cognitive impairment, depressive symptoms, and the overall level of functioning.

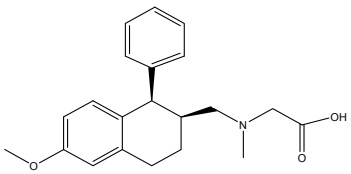


Figure 3.1
Chemical structure of Org 25935

METHODS

GIANT was a multicenter, 12-week double-blind, parallel-group, randomized clinical trial of adjunctive Org 25935 or placebo for the treatment of predominant persistent negative symptoms of schizophrenia. It was conducted according to Good Clinical Practice guidelines at 25 sites across Europe (Finland, Norway, France, the Czech Republic), Russia, and Latin America (Argentina, Chile) from April 2007 through September 2008. The trial protocol was approved by the independent ethics committee at each site, and all subjects provided written informed consent after the scope and nature of the investigation, including recording of interviews for second opinion, had been explained to them before screening.

Subjects

Subjects of both sexes meeting Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) criteria for schizophrenia (nonfirst episode) with predominant persistent negative symptoms were eligible for the study, if they were in the age range of 18 and 55 years of age, receiving stabilized treatment for ≥3 months with an SGA other than clozapine (without any change in psychiatric care or SGA dosing regimen during ≥4 weeks prior to screening), and continued to experience predominant negative symptoms, presenting: a score ≥4 on three or more of the following core items of the Positive and Negative Syndrome Scale (PANSS)[14] at screening: blunted affect (N1), emotional withdrawal (N2), poor rapport (N3), passive social withdrawal (N4), lack of spontaneity (N6), motor retardation (G7), active social avoidance (G16); an overall summary score >20 on these core items; a score ≥5 ("marked" severity or higher) on less than two of following PANSS items: delusions [P1], hallucinatory behavior [P3], excitement [P4], grandiosity [P5], or suspiciousness / persecution [P6]; a score <20 on the PANSS positive subscale; a score <9 on the Calgary Depression Scale for Schizophrenia (CDSS)[15]; a score ≤3 on the clinical global impression of Parkinsonism of the adapted version of the Extrapyramidal Symptom Rating Scale (ESRS)[16] at screening.

Subjects were also required to have a caregiver or identified responsible person to support the compliance of the subject with the study procedures, and to be medically stable (with stable drug treatment ≥4 weeks for any medical condition). Furthermore, subjects were not to be at imminent risk of self-harm as documented by a score ≤9 on the InterSePT Scale for Suicidal Thinking (ISST).[17]

Key exclusion criteria were a history of any seizure disorder or progressive eye disease, treatment with clozapine or any investigational drug, a diagnosis of alcohol or drug dependence within six months before screening, history of a malignancy or other chronic and/or degenerative processes, as well as an imminent risk to harm others.

Treatment

Eligible subjects were randomly assigned (in a 1:1:1 ratio) to treatment with Org 25935 4–8 mg twice daily ("low Org"), Org 25935 12–16 mg twice daily ("high Org"), or matching placebo in addition to the continued treatment with an SGA. The actual dose at which the subject was treated from study day 7 onwards was at the discretion of the investigator. During

the first six weeks, truncation of upward titration to the higher dose and, if necessary, downward titration to the starting dose were allowed. After that, subjects were to be maintained on the dose to which they had been titrated for the remainder of the trial period. Background SGA treatment had to be continued unchanged during the study.

Subjects were allowed concomitant use of benzodiazepines at a dose equivalent to maximally 4 mg lorazepam per day, and could remain on any anticholinergic, beta blocker, antidepressant, anti-anxiety, or hypnotic treatment that had been started ≥4 weeks before screening.

The investigational product (IP) was administered as oral tablets, provided in blisters. Formulation of Org 25935 and matching placebo was indistinguishable with respect to appearance, shape, smell, and taste. Adherence to treatment was assessed by medication review and tablet count at all visits. Subjects were to be withdrawn from the study whenever compliance became less than 70% over the entire trial period (from baseline onwards), either for the SGA or IP.

A blood sample for assessment of Org 25935 concentration was taken during the visit on days 14, 42 and 84 (Endpoint) at a random time point after last dose taken. The time and date of blood sampling for assessment of each sample together with the time of dosing of trial medication was recorded on the lab requisition form. Bioanalysis was performed according to a validated assay under Good Laboratory Practice (GLP). In order to keep personnel blinded to the treatment allocation of individual subjects, samples were relabeled at the central laboratory with virtual subject numbers. After unblinding of the trial, mean plasma levels were listed by treatment group.

Clinical Assessments

Overall severity of negative symptoms was measured using the composite score on the first 22 items of the Scale for Assessment of Negative Symptoms ($SANS_{1-22}$).[18] The $SANS_{1-22}$ composite score was chosen as the primary efficacy variable based on the majority of expert recommendations from a dedicated advisory board held in August 2005 (unpublished). SANS assessments were conducted at baseline, week 3, week 6, and week 12 (endpoint).

Alternative measures of negative symptoms included the SANS summary score (sum score of the five global items 7, 13, 17, 22, and 25), PANSS negative subscale, and PANSS Marder factor score for negative symptoms (PANSS-M).[19] The PANSS positive subscale, CDSS, and

ESRS were used to assess positive, depressive, and extrapyramidal symptoms, respectively. PANSS, CDSS, and ESRS assessments were done at screening, week 4, week 8, and endpoint. The scale for Global Assessment of Functioning (GAF)[20] was used to assess global severity at all visits (screening, week 1, 2, 3, 4, 6, 8, 10, and endpoint). The Neurological Evaluation Scale (NES)[21] was used to categorize subjects into those with a baseline level of neurological soft signs below (NES <5) or above (NES ≥5) the median, and to measure change in these symptoms at endpoint.

Vital signs and body weight were assessed at each visit. Physical examinations were carried out at screening and endpoint. Electrocardiograms (ECTs) were performed at screening, week 3, and endpoint. Visual function, including measurements of acuity, visual field, and color/hue vision, was evaluated at screening and days 21 and 70, when facilities were locally available (in 14 of the 25 sites). Visual acuity was measured using ETDRS charts[22], possible defects in the visual field were determined using automated perimetry (Humphrey 30–2 or equivalent)[23], and color/hue perception was tested by determining the Bowman score with a computerized version of the Farnsworth Dichotomous test (D-15).[24] Routine blood and urine tests were undertaken at screening, week 2, week 6, and endpoint. Tolerability was assessed in terms of (serious) adverse events ([S]AEs), either reported by the subject or observed by the investigator at any visit, up to 7 days after the last dose (30 days for SAEs).

Rater Training and Qualification

Raters on SANS, PANSS, and GAF were required to have at least two years' experience with the clinical evaluation of schizophrenia, and have a medical degree (MD) to qualify as a potential rater for GAF, ESRS, NES, and ISST. Rater candidates were trained on patient interviewing skills, and needed to qualify as a rater for SANS and PANSS by independently viewing, scoring, and submitting scores for a standard set of audiovisual interview recordings. Candidates were admitted as a rater for SANS when their scores did not deviate more than ±1 from the Gold Standard Rating (GSR) on the five global sub-scores of SANS, and $SANS_{1\text{-}22}$ showed less than 15% deviation from GSR. Candidates were admitted as a rater for PANSS when their scores for at least 80% of the items were within the predefined GSR range.

In this trial, a new method aimed at improving the quality of psychopathological ratings at the participating sites by providing timely feed-back from independent expert raters was

tested for the first time.[25,26] This so-called Expert Rater Assisted Score Evaluation (ERASE) method will be described in more detail elsewhere. The results reported in this manuscript are based on the final ratings of the site raters, after having been allowed to make adjustments to their preliminary system scores on the basis of input from independent raters. Explicit informed consent was obtained from all patients allowing the implementation of a video monitoring system, whereby independent raters, native speaker of the local language and blind to preliminary site rater scores, reviewed audiovisual recordings of patient interviews for PANS (at screening) and SANS (at baseline and endpoint), did an independent severity grading of symptoms, and subsequently shared their scores with the site raters.

Cognitive Battery

The cognitive domains of verbal memory, visual memory, speed of processing, perception of emotions (social cognition), reasoning, executive functioning (problem solving), working memory, and sustained attention were assessed using the computerized CNS-Vital Signs Neurocognitive Test Battery at baseline, week 6, and endpoint. The Neuro Cognitive Index (NCI), based on a weighted calculation of individual item scores, was used as an overall measure of cognitive functioning. The cognitive battery included tests for verbal memory (repeats), visual memory (repeats), symbol digit coding, shifting attention, emotional acuity, reasoning, and four part continuous performance (also known as the "N-Back" CPT).

Statistical Analyses

Sample size calculation indicated that 60 subjects per treatment arm were sufficient to demonstrate an effect size (ES) of 0.50 at a power of 80%. This ES shows a meaningful improvement of 4.5 points over placebo in $SANS_{1-22}$ composite score change from baseline, assuming a standard deviation of 9.0. Because sample size calculation did not compensate for drop-out rates (assuming minimum treatment duration of eight weeks is necessary to measure improvement in negative symptoms), subjects dropping out before week 8 were to be replaced.

The efficacy analysis for the primary and secondary variables was performed for all subjects in the Intent-To-Treat (ITT) population who had been assessed for efficacy after at least eight weeks treatment (ITT-Wk8). The approach was taken because this was a proof-of-concept trial in a limited number of subjects, whereby a sufficiently long exposure to drug

(close to the total duration as by protocol) was deemed necessary to demonstrate any meaningful change in behavior and emotions over time.

For all continuous variables, change from baseline was analyzed using an analysis of covariance (ANCOVA) model with the baseline value as covariate and with treatment group and center as fixed factors. Treatment differences between each of the two Org 25935 groups and placebo, as well as 95% confidence intervals of the treatment differences, including P-values, based on the ANCOVA model, were calculated. Treatment by center interaction was explored. Model assumptions (normality and homoscedasticity) were checked by visual inspection of the residuals. The statistical test of each Org 25935 group against placebo was performed two-sided at the 0.05 significance level, using a last-observation-carried-forward (LOCF) method to impute missing data. No multiplicity correction was made.

A frequency table was prepared for occurring (S)AEs by treatment and preferred term, but no statistical tests were performed for safety data.

A path analytic approach was taken to explore whether a differential efficacy for negative symptoms favoring either Org 25935 or placebo was likely to be due to direct and/or indirect therapeutic effects.[27] Assuming a linear relationship between negative symptoms and covariates in other symptom domains, the direct effect on negative symptoms was estimated using two regression analyses. One according to an extended model and the second according to a reduced model, where the resulting ratio in the average "rate of change" (b_2/b_1) provided an estimate of the degree by which the treatment effect on negative symptoms ($SANS_{1-22}$ change from baseline) could not be explained by concurrent changes in positive (PANSS+=PANSS positive subscale), cognitive (NCI), depressive (CDSS), and extrapyramidal (ESRS) symptoms:

1. A regression analysis with reduced model:

$$SANS_{1-22} = b_1*\text{Treatment}$$

2. A regression analysis with extended model:

$$SANS_{1-22} = b_2*\text{Treatment} + b_p*\text{PANSS+} + b_e*\text{ESRS} + b_c*\text{NCI} + b_d*\text{CDSS}$$

Data in Org 25935 groups were pooled for these multiple regression analyses.

RESULTS

Subject Characterization

Drop-out rates throughout the study were low and similar across treatment arms: low Org (15%), high Org (11%), and placebo (13%). Baseline characteristics by treatment, including background treatment, were comparable without statistically significant differences between treatment groups (Table 3.1; P-values not shown).

Treatment Adherence

Patient compliance with dosing instructions as measured by pill count was comparably high in all three treatment groups and was on average 99% with intake of prescribed antipsychotics and 97% with IP intake. Among subjects exposed to Org, the mean daily dose was 10.9±2.9 mg in the low Org group (28 subjects (19%) receiving a final dose of 4 mg BID and 43 subjects (30%) receiving a final dose of 8 mg BID), and 25.8±2.9 mg in the high Org group (27 subjects (19%) receiving a final dose of 12 mg BID and 46 subjects (32%) receiving a final dose of 16 mg BID).

Pharmacokinetic Results

Mean (8.4-10.5 hrs) and median (10.0-11.9 hrs) sampling time after last drug intake were comparable in both active dose groups. Mean plasma level at days 14, 42, and 84 in the low Org group was 83, 108, and 128 ng/mL respectively, and in the high Org group 173, 229, and 217 ng/mL. Mean plasma level by received dose was 81, 65, and 72 ng/mL on 4 mg BID, 86, 126, and 159 ng/mL on 8 mg BID, 188, 186, and 165 on 12 mg BID, and 152, 250, and 243 ng/mL on 16 mg BID respectively.

Negative Symptoms

When analyzing the primary efficacy outcome parameter, subjects in the ITT-Wk8 population showed a mean reduction of 17% to 24% in negative symptoms ($SANS_{1-22}$) without statistically significant differences between low Org or high Org in comparison with placebo (Table 3.2, Figure 3.2). The effect size (Cohen's D) on low Org was –0.19 and on high Org 0.03. Similar results were obtained when comparing PANSS negative subscale (Table 3.2, Figure 3.3*a*) and the PANSS-M (Table 3.2) scores, the low Org group showing a trend

towards higher reduction in scores versus placebo without reaching statistically significant difference at any time point. When defining response as an individual improvement of at least 20% as compared with baseline, no significant differences were observed for the PANSS-M nor the $SANS_{1-22}$ scores, though also in this post hoc analysis the low Org group had a numerical advantage over placebo (P=0.06).

Table 3.1

Baseline subject characteristics

	Org 25935 4–8 mg bid (N = 71)	Org 25935 12–16 mg bid (N = 73)	Placebo (N = 70)
Age, mean ± SD	37.4 ± 9.5	38.8 ± 11.0	38.1 ± 10.5
Gender, female, n (%)	30 (42.3)	27 (37.0)	24 (34.3)
Smoking, Yes, n (%)	32 (45.1)	31 (42.5)	26 (37.1)
Marital status, married, n (%)	11 (15.5)	15 (20.5)	5 (7.1)
Negative symptoms persistence, n (%)			
>5 years	31 (43.7)	31 (42.5)	37 (52.9)
Neurological soft signs (NES), n (%)			
Below median (<5)	31 (44.3)	31 (43.1)	38 (55.1)
SANS, mean ± SD			
$SANS_{1-22}$ composite score	60.1 ± 10.7	63.4 ± 10.1	63.9 ± 11.9
Blunted affect	2.9 ± 0.8	3.3 ± 0.6	3.2 ± 0.9
Alogia	2.3 ± 0.9	2.6 ± 0.8	2.5 ± 0.8
Avolition	3.3 ± 0.8	3.2 ± 0.8	3.4 ± 0.8
Anhedonia	3.7 ± 0.7	3.6 ± 0.7	3.9 ± 0.6
Attention	2.3 ± 1.0	2.3 ± 1.0	2.3 ± 0.8
PANSS, mean ± SD			
Total score	77.6 ± 11.5	80.9 ± 11.5	79.3 ± 10.4
Positive subscale	12.9 ± 3.3	13.3 ± 3.1	13.4 ± 3.0
Negative subscale	26.3 ± 3.7	27.1 ± 3.8	26.7 ± 3.9
General subscale	38.4 ± 7.1	40.5 ± 7.4	39.1 ± 6.7
Negative Marder factor	26.6 ± 3.6	27.4 ± 3.7	26.8 ± 3.9
GAF, mean ± SD	51.9 ± 10.6	51.3 ± 9.1	50.1 ± 10.6
NES, mean ± SD	6.2 ± 5.0	5.5 ± 4.9	5.8 ± 5.1
CDSS, mean ± SD	2.4 ± 2.4	2.5 ± 2.3	2.6 ± 2.5
ESRS, mean ± SD	2.4 ± 4.5	2.1 ± 4.3	2.4 ± 4.3
Neurocognitive Index (NCI), mean ± SD	77.9 ± 15.4	75.6 ± 14.2	75.6 ± 14.3

In further sub-analysis stratifying by baseline neurological soft signs, in subjects with relatively mild soft signs (NES <5 at baseline), mean ± sd $SANS_{1-22}$ change from baseline was –16.22 ± 14.69 on low Org (n=27), –11.07 ± 14.17 on high Org (n=28), and –10.79 ± 10.88 on placebo (n=33). The difference between low Org and placebo was statistically significant (P=0.02), but not between high Org and placebo (P=0.62).

Table 3.2

Clinical improvement from Baseline to Endpoint in individual symptom domains

	Org 25935 4–8 mg Twice-daily (N = 62)				Org 25935 12–16 mg Twice-daily (N = 67)				Placebo (N = 62)		
	CFB	SD	%	*P**	CFB	SD	%	*P**	CFB	SD	%
SANS											
Composite score $_{1-22}$	–13.50	12.58	–23.8	0.27	–10.91	11.85	–17.1	0.97	–11.21	11.28	–17.8
Sum (global items)	–3.13	3.29	–22.0	0.35	–2.78	3.08	–18.3	0.68	–2.65	2.87	–17.4
Blunted affect	–0.56	0.76	–21.4	0.66	–0.63	0.76	–18.9	0.61	–0.58	0.71	–18.5
Alogia	–0.56	0.92	–25.7	0.34	–0.49	0.84	–17.0	0.77	–0.50	0.67	–21.1
Avolition	–0.85	1.01	–26.0	0.08	–0.67	0.77	–20.4	0.52	–0.61	0.88	–17.9
Anhedonia	–0.68	0.78	–18.0	0.13	–0.48	0.93	–11.7	0.67	–0.52	0.76	–11.7
Attention	–0.47	1.08	–10.9	0.94	–0.51	1.08	–15.8	0.78	–0.44	0.88	–16.8
PANSS											
Total score	–11.84	10.34	–25.8	0.11	–10.69	10.40	–20.9	0.64	–9.73	11.23	–19.3
Positive subscale	–1.44	1.93	–21.7	0.19	–1.61	4.02	–23.2	0.14	–1.00	2.69	–18.8
Negative subscale	–4.77	4.74	–25.3	0.22	–3.64	3.94	–17.8	0.70	–3.95	3.94	–20.1
General subscale	5.63	5.80	–25.4	0.13	–5.43	5.21	–21.8	0.85	–4.77	6.16	–17.2
Negative Marder Factor	–5.44	5.26	–29.3	0.08	–4.40	4.18	–22.1	0.88	–4.26	4.11	–22.6
GAF	7.85	9.00	16.3	0.39	6.69	7.18	13.6	0.95	7.10	7.78	15.7
CDSS	–0.95	1.88	–48.6	0.04	–0.76	1.88	–37.0	0.16	–0.56	2.67	–26.3
ESRS	–0.80	2.20	–25.9	0.66	–0.30	1.20	–17.2	0.26	–1.0	3.40	–32.0
Cognitive battery											
NCI	3.21	9.99	4.5	0.50	4.14	10.74	6.7	0.38	2.32	14.20	3.5
Verbal memory	–0.76	7.61	–0.2	0.27	0.41	7.85	3.0	0.22	–0.78	8.60	0.1
Visual memory	–0.35	4.98	–0.1	0.10	–1.73	6.20	–3.0	0.82	–1.34	6.59	–2.1
Sustained attention	0.96	10.76	–35.8	0.74	3.81	11.14	24.9	0.55	1.95	15.88	71.5
Working memory	0.96	4.62	–0.4	0.44	0.37	4.75	4.0	0.49	0.82	6.72	–37.9
Executive functioning	10.16	19.55	34.2	0.99	9.33	20.09	39.7	0.73	12.57	23.26	43.8
Reasoning	0.65	4.30	–38.2	0.84	0.66	3.95	28.8	0.82	0.78	4.68	23.4
Speed of processing	5.62	8.53	29.7	0.63	5.73	13.21	–9.5	0.89	5.17	12.94	48.0
Memory score	–1.26	10.64	–0.9	0.15	–1.32	11.61	–0.3	0.36	–2.12	12.35	–1.6
Perception emotions	1.42	4.18	3.5	0.02	1.55	4.35	16.1	0.02	–0.80	5.91	–44.5

* Differences in change from baseline between the 2 doses of Org 25935 versus placebo were pair-wise tested with an ANCOVA model with baseline as a covariate and pooled sites as a factor, using LOCF approach.

CFB = change from baseline, SD = standard deviation.

Analyses of potential "pseudospecificity" of effects on negative symptoms showed that the rate of change in $SANS_{1-22}$ was –0.017 according to the reduced model and –0.071 according to the extended model. This means that according to the used model, 1 - (0.017/0.071) = (1-0.24)*100 or 76% of reduction in $SANS_{1-22}$ can be explained by concurrent changes in

positive, general psychopathological, depressive, neurocognitive, or extrapyramidal symptoms.

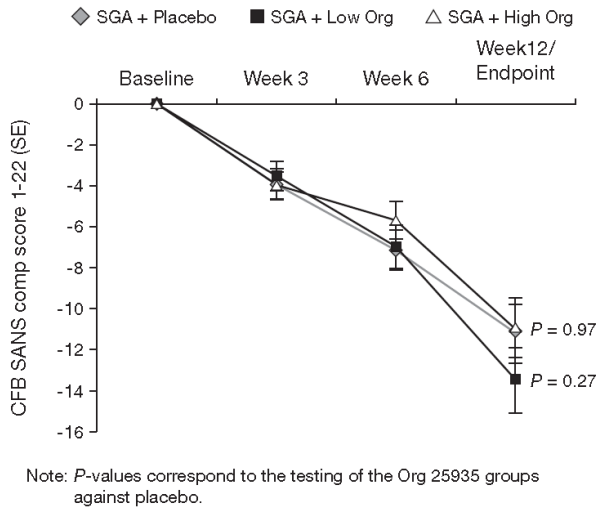


Figure 3.2

Mean change from Baseline in $SANS_{1-22}$ composite score by adjunctive treatment in subjects having completed at least eight weeks double-blind treatment, using the LOCF approach

Additional Effects

Improvement in additional secondary efficacy variables is presented in Table 3.2. The relative change from baseline (in %) in SANS summary score and PANSS (total and subscales) are of similar magnitude (18%–26%) as the observed 17% to 24% improvement in $SANS_{1-22}$, without significant differences between treatment groups.

The results for the score changes on the PANSS total score (Figure 3.3*b*), negative and positive subscales (Figures 3.3*a* and 3.3*c* respectively), GAF, CDSS, and tested cognitive domains are summarized in Table 3.2. Note that subjects with too many positive symptoms were to be excluded from the study.

Generally, no statistically significant differences were observed for either dose group of Org 25935 compared with placebo with a few exceptions. Depressive symptoms as measured by CDSS and the perception of emotions test (POET) as part of the CNS-VS battery showed descriptively a statistically significant improvement (not corrected for multiplicity) over

placebo in the low dose Org 25925 group at an effect size of –0.17 and 0.43, respectively. POET showed also statistically significant improvement over placebo (effect size = 0.46) in the high Org 25935 dose group.

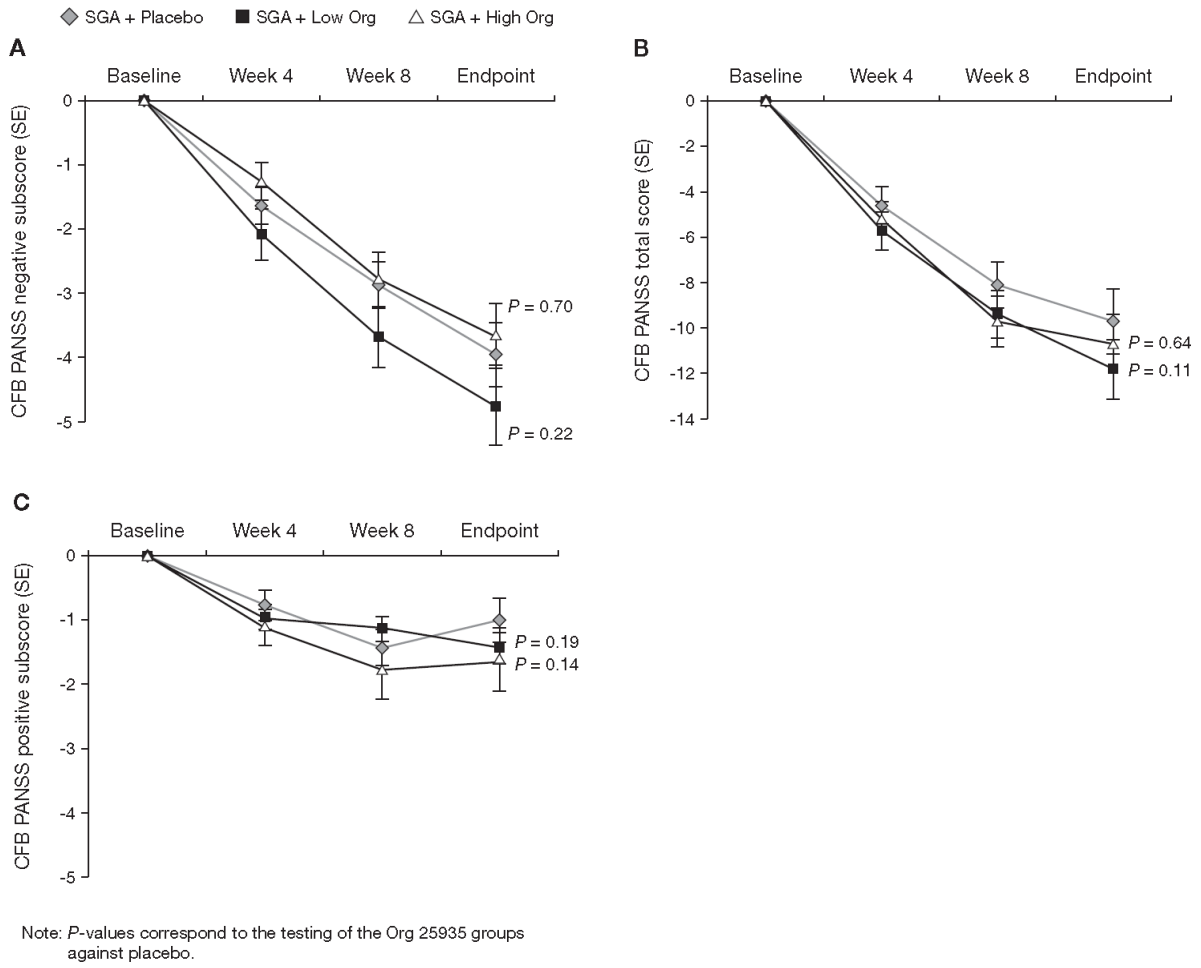


Figure 3.3

Mean change from baseline (CFB) in PANSS negative subscale score (*a*), PANSS total score (*b*), and PANSS positive subscale score (*c*) at secondary efficacy end point, using the LOCF approach.

Adverse Events

Treatment with Org 25935 appeared to be generally well tolerated and safe. A total of five serious adverse events were reported during the trial (of which three were related to worsening of the underlying disease, one due to dizziness, one subject experienced a convulsion, and one subject had a foot fracture after accidental collapse). There were no AEs

occurring at a frequency higher than 10% in any of the treatment groups, and most AEs were mild in intensity. AEs of which the incidence appeared to be higher on Org 25935 than placebo and to increase with dose, included (percentages for subjects on placebo, low Org, and high Org, respectively): dizziness (0.0%, 2.8%, 9.6%), nasopharyngitis (1.4%, 4.2%, 9.6%), anxiety (2.9%, 2.8%, 5.5%), somnolence (1.4%, 2.8%, 5.5%), and blurred vision (0.0%, 1.4%, 4.1%).

No increase in EPS was found, as assessed by the ESRS scale (Table 3.2). No effects were seen on any safety parameters, including routine laboratory panels, ECGs, vital signs, and body weight that can be interpreted as clinically meaningful.

Ophthalmologic assessments were done in 66 subjects. Sporadic minor changes in visual field and acuity were observed. However, the changes were not considered clinically relevant since none of the affected subjects experienced a visual AE. A clinically significant abnormality in visual field was determined in two subjects, of which one subject was on placebo.

DISCUSSION

This study was to our knowledge the first dedicated clinical trial with a selective and potent GlyT-1 inhibitor in subjects with predominant persistent negative symptoms in schizophrenia. The trial did not provide evidence of significant clinical benefits for either dose group of Org 25935 on negative symptoms as measured by the prespecified, primary outcome variable. When using data generated without the input from independent raters, ie, the preliminary, unadjusted site rater scores on SANS, the overall trial outcome was the same. Over a period of 12 weeks, a substantial reduction of symptoms was seen with adjunctive placebo treatment, which did not differ significantly on most measures from the adjunctive treatment with either dose regimen of GlyT-1 inhibitor. Although in some sub-analyses numerical differences were noted, generally trending in favor of the low Org group, sensitivity analyses generally did not support the presence of a robust signal in this trial. Also measures of overall functional improvement and effect sizes did not point towards clinically relevant difference between treatment groups. There may have been a difference between the subgroup of subjects with a relatively low level of neurological soft signs (NSS) and higher levels of NSS at baseline.

Presence of NSS has earlier been associated with negative symptoms and poor clinical response.[28-30] The clinical relevance of a positive outcome in this subpopulation is, however, disputable, since "pseudospecificity" analyses do not appear to support the conclusion of convincing drug effects on negative symptoms in this trial.

We also did not find evidence for a beneficial effect of either dose of adjunctive Org 25935 on most cognitive domains as measured by the CNS-VS computerized battery. A possible exception is the result on the perception of emotions, as measured by the POET test, which might be subject for further explorations.

A potential limitation of the study was the relatively high improvement (~18% reduction in SANS composite score) and low drop-out rate (13% over 12 weeks) on placebo compared with results obtained in earlier adjunctive, negative symptom trials. During the first 12 weeks of a 6-month study with adjunctive D-cycloserine in subjects with prominent negative symptoms, a drop-out rate of 10/28 (36%) and a negative symptom improvement of 3.4% was observed on placebo.[31] In an 8-week trial with glycine added to clozapine, both drop-out and negative symptom improvement rates on placebo were very small.[32] In an 8-week trial with pregnenolone added to an SGA, no drop-outs and a 5% reduction of negative symptom severity were observed on placebo.[33] In a 6-week study with mirtazapine added to poorly responsive subjects on first-generation antipsychotics, the drop-out rate was very low and the negative symptom severity improved with 3% on placebo.[34] Also due to the lack of an established active positive control in this indication, it remains unclear at this stage to what extent the relatively high placebo response in the GIANT trial may have contributed to the lack of significant findings with Org 25935, neither what may have caused the favorable response on placebo. Since most patients with schizophrenia have mixed positive and negative symptoms, the selection of patients with predominant negative symptoms can be considered a biased sampling. The study might have been less vulnerable to non-specific treatment effects when patients with high positive symptoms would not have been excluded from participation.

We would like to add a comment on the dose design at 4-8 and 12-16 mg Org 25935. Two non-overlapping dose regimens were selected within the predicted therapeutic range, based on PK/PD results obtained in healthy volunteers, whereby a steady increase in CSF glycine level was observed after administration of single rising dosages of 4, 8 and 16 mg Org 25935 and disturbing, visual adverse events were considered prohibitive for further dose increments

beyond a twice daily 16 mg level. The overall results in the current study may seem in contrast to recently announced results with another selective inhibitor of GlyT-1 (RG 1678), for which superiority over placebo was claimed in the lower dose range in a similar treatment setting.[35] At closer inspection, however, the results of both studies demonstrate some similarities. In view of the fact that in the respective study the positive results were seen at lower doses and potential signals of efficacy in the GIANT trial occurred largely in the low-dose group, it cannot be excluded that the drug Org 25935 has been administered in too high doses for a clinical benefit in negative symptoms. PET scan studies with Org 25935 might provide insight into GlyT-1 binding at various dose levels and possible saturation thresholds, before clinical studies with lower dose levels are undertaken to further investigate this issue. Although we have no comparable data on Org 25935 that would allow predicting equivalence to doses of RG1678 that were reported to be effective, both 16 mg Org 25935 and 30 mg/kg RG1678 have been shown to increase glycine CSF plasma levels up to 2.5 times baseline levels in humans and rats respectively.[13,36] On the other hand, an important differentiating characteristic of the GlyT-1 inhibitors is the way these bind to the transporter complex. RG1678 binds non-competitively, whereas Org 25935 can be expected to bind competitively to the transporter as a result of a N-methyl-glycine moiety that is present in its molecular structure, possibly rendering the binding of Org 25935 as being sensitive to glycine concentrations.[10] Both Org 25935 and RG1678 are highly selective inhibitors of GlyT-1, and further studies will be required to investigate whether the mode of interaction with the transporter complex can have implications for the potential clinical usefulness of specific GlyT-1 inhibitors.

ACKNOWLEDGMENTS

GIANT Study Group: Clínica Privada Neuropsiquiátrica San Agustín, La Plata (Argentina): A. Bertoldi, M.D., Ph.D.; FLENI, Buenos Aires (Argentina): R. Fahrer, M.D., Ph.D.; Clínica Psiquiátrica San Jorge, Lanús Este (Argentina): A. Mega, M.D., Ph.D.; Clínica Privada Banfield, Banfield (Argentina): E. Amado Cattaneo, M.D.; Directora Instituto Neurociencias, Mendoza (Argentina): R. Galeno, M.D.; Universitätsklinikum Innsbruck (Austria): A. Hofer,

M.D.; Private practice, Salzburg (Austria): A. Whitworth, M.D.; Cipam – Clínica Pedro Montt, Santiago (Chile): V. Larach, M.D., Ph.D.; Centro de Estudios y Tratamiento de Enfermedades Psiquiátricas, Santiago (Chile): G. Vergara, M.D.; Hospital Base de Valdivia, Valdivia (Chile): F. Bertran, M.D.; Nestátni zdravotnické zarizeni, Melnik (Czech Republic): H. Matlakova, M.D.; Terapie.info, Prague (Czech Republic): T. Rektor, M.D.; Telemens, Prerov (Czech Republic): J. Rektor, M.D.; Psychiatricka ambulance, Kutna Hora (Czech Republic): S. Pietrucha, M.D.; Private clinic, Brno (Czech Republic): R. Prikryl; Private clinic, Kuopio (Finland): J. Jääskeläinen, M.D.; Paloniemen sairaala, Hormajärvi (Finland): H. Katila, M.D.; Niuvanniemen sairaala, Kuopio (Finland): J. Tiihonen, M.D., Ph.D.; Kellokosken sairaala, Kellokoski (Finland): E. Sailas, M.D.; Helsinki University Central Hospital, Helsinki (Finland): G. Joffe, M.D., Ph.D.; CHS La Chartreus – Dijon Cedex (France): G. Badet, M.D.; CHU de Grenoble – Hôpital Sud, Grenoble Cedex (France): T. Bougerol, M.D., Ph.D.; Centre Medico Psychologique, Dole (France): B. Bonnaffoux, M.D.; Centre Médical Claude Bernard, Elancourt (France): J. Gailledreau, M.D.; Brattvåg Legekontor, Brattvåg (Norway): O.J. Høyberg, M.D.; Jæren D.P.S., Bryne (Norway): S. Heskestad, M.D., Ph.D.; Private practice, Oslo (Norway): Ø. Kavlie, M.D.; City Psychiatric Hospital, Saratov (Russia): V. Vilyanov, M.D., Ph.D.; Psychoneurological Hospital (Smolensk State Medical Academy), Smolensk (Russia): A. Okhapkin, M.D., Ph.D.; Samara Psychiatric Hospital, Samara (Russia): M. Sheyfer, M.D.; Treatment and Rehabilitation Research Center “Phoenix”, Rostov-on-Don (Russia): A. Bukhanovsky, M.D., Ph.D.; Psychoneurology Health Center, Saint-Petersburg (Russia): V. Tadtayev, M.D.; Psychiatric Hospital, Saint-Petersburg (Russia): M. Popov, M.D.

REFERENCES

1. Buchanan RW. Persistent negative symptoms in schizophrenia: an overview. *Schizophrenia Bulletin* 2007;33(4):1013-1022.
2. Tamminga CA. The neurobiology of cognition in schizophrenia. *Journal of Clinical Psychiatry* 2006;67:9-13.
3. Insel TR. Rethinking schizophrenia. *Nature* 2010;468(7321):187-193.
4. Goff DC, Coyle JT. The emerging role of glutamate in the pathophysiology and treatment of schizophrenia. *American Journal of Psychiatry* 2001;158:1367-1377.
5. Tsai GE, Lin P-Y. Strategies to enhance N-methyl-D-aspartate receptor-mediated neurotransmission in schizophrenia, a critical review and meta-analysis. *Current Pharmaceutical Design* 2010;16(5):522-537.
6. Javitt DC. Glycine transport inhibitors and the treatment of schizophrenia. *Biological Psychiatry* 2008;63(1):6-8.
7. Javitt DC. Glycine transport inhibitors for the treatment of schizophrenia: symptom and disease modification. *Current Opinion in Drug Discovery & Development* 2009;12(4):468-478.
8. Sur C, Kinney GG. The therapeutic potential of glycine transporter-1 inhibitors. *Expert Opinion on Investigational Drugs* 2004;13(5):515-521.
9. Walker G, Hamilton W, Ge J, Bruin J, Shahid M, Hill D. ORG 25935: a selective glycine uptake inhibitor. *Schizophrenia Research* 2001;49(1-2):97-97.
10. Mezler M, Hornberger W, Mueller R, Schmidt M, Amberg W, Braje W, Ochse M, Schoemaker H, Behl B. Inhibitors of GlyT1 affect glycine transport via discrete binding sites. *Molecular Pharmacology* 2008;74(6):1705-1715.
11. Ge J, Hamilton W, Shahid M, Hill D, Walker G. The effects of ORG 25935 on the extracellular levels of glycine in brain regions of freely moving rats. *British Journal of Pharmacology* 2001;133:135.
12. Lidö HH, Stomberg R, Fagerberg A, Ericson M, Söderpalm B. The Glycine Reuptake Inhibitor Org 25935 Interacts With Basal and Ethanol-Induced Dopamine Release in Rat Nucleus Accumbens. *Alcoholism: Clinical and Experimental Research* 2009;33(7):1151-1157.
13. Kleijn H, Gerrits M, Schoemaker J, de Greeff H. PK-PD Modelling and simulation of phase I data to support development of a new compound in the field of CNS. *5th International Symposium on Measurement and Kinetics of In Vivo Drug Effects,* April 26–29, Noordwijkerhout, The Netherlands 2006.
14. Kay SR, Flszbein A, Opfer LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 1987;13(2):261-276.
15. Addington D, Addington J, Maticka-Tyndale E. Assessing depression in schizophrenia: the Calgary Depression Scale. *British Journal of Psychiatry* 1993.
16. Chouinard G, Margolese HC. Manual for the extrapyramidal symptom rating scale (ESRS). *Schizophrenia Research* 2005;76(2):247-265.
17. Lindenmayer J, Czobor P, Alphs L, Nathan A-M, Anand R, Islam Z, Chou JC, Group IS. The InterSePT scale for suicidal thinking reliability and validity. *Schizophrenia Research* 2003;63(1):161-170.
18. Andreasen NC. Negative symptoms in schizophrenia: definition and reliability. *Archives of General Psychiatry* 1982;39(7):784-788.
19. Marder SR, Chouinard G. The effects of risperidone on the five dimensions of schizophrenia derived by factor analysis: combined results of the North American trials. *Journal of Clinical Psychiatry* 1997;58(12):1,478-546.
20. First MB. Diagnostic and statistical manual of mental disorders. *DSM IV-4th edition. APA* 1994:1994.
21. Buchanan RW, Heinrichs DW. The Neurological Evaluation Scale (NES): a structured instrument for the assessment of neurological signs in schizophrenia. *Psychiatry Research* 1989;27(3):335-350.
22. Ferris FL, Kassoff A, Bresnick GH, Bailey I. New visual acuity charts for clinical research. *American Journal of Ophthalmology* 1982;94(1):91-96.
23. Beck RW, Bergstrom TJ, Lighter PR. A clinical comparison of visual field testing with a new automated perimeter, the Humphrey Field Analyzer, and the Goldmann perimeter. *Ophthalmology* 1985;92(1):77-82.
24. Bowman K. A method for quantitative scoring of the Farnsworth Panel D-15. *Acta Ophthalmologica* 1982;60(6):907-916.
25. Schoemaker J, Gaur R, Chawla V, Jansen W, Szegedi A. Expert rater assisted score evaluation (ERASE): a new method to enhance signal detection in randomized, placebo-controlled, clinical trials.

Paper presented at: *International Society for CNS Trials Methodology 4th Annual Meeting*, Washington DC, USA. 2009.
26. Schoemaker JH, Inventor; Google Patents, assignee. Accurate method to assess disease severity in clinical trials concerning psychopathology. 2008.
27. Möller H-J, Müller H, Borison RL, Schooler NR, Chouinard G. A path-analytical approach to differentiate between direct and indirect drug effects on negative symptoms in schizophrenic patients. *European Archives of Psychiatry and Clinical Neuroscience* 1995;245(1):45-49.
28. Merriam AE, Kay SR, Opler LA, Kushner SF, Van Praag HM. Neurological signs and the positive-negative dimension in schizophrenia. *Biological Psychiatry* 1990;28(3):181-192.
29. Smith RC, Kadewari RP. Neurological soft signs and response to risperidone in chronic schizophrenia. *Biological Psychiatry* 1996;40(10):1056-1059.
30. Smith RC, Kadewari RP, Rosenberger JR, Bhattacharyya A. Nonresponding schizophrenia: differentiation by neurological soft signs and neuropsychological tests. *Schizophrenia Bulletin* 1999;25(4):813-825.
31. Goff DC, Herz L, Posever T, Shih V, Tsai G, Henderson DC, Freudenreich O, Evins AE, Yovel I, Zhang H, Schoenfeld D. A six-month, placebo-controlled trial of D-cycloserine co-administered with conventional antipsychotics in schizophrenia patients. *Psychopharmacology* 2005;179(1):144-150.
32. Evins AE, Fitzgerald SM, Wine L, Rosselli R, Goff DC. Placebo-controlled trial of glycine added to clozapine in schizophrenia. *American Journal of Psychiatry* 2000;157:826-828.
33. Marx CE, Keefe RS, Buchanan RW, Hamer RM, Kilts JD, Bradford DW, Strauss JL, Naylor JC, Payne VM, Lieberman JA, Savitz AJ, Leimone LA, Dunn L, Porcu P, Morrow AL, Shampine LJ. Proof-of-concept trial with the neurosteroid pregnenolone targeting cognitive and negative symptoms in schizophrenia. *Neuropsychopharmacology* 2009;34(8):1885-1903.
34. Joffe G, Terevnikov V, Joffe M, Stenberg JH, Burkin M, Tiihonen J. Add-on mirtazapine enhances antipsychotic effect of first generation antipsychotics in schizophrenia: a double-blind, randomized, placebo-controlled trial. *Schizophrenia Research* 2009;108(1-3):245-251.
35. Umbricht D, Yoo K, Youssef E, Dorflinger E, Martin-Facklam M, Bausch A. Investigational glycine transporter type 1 (GlyT1) inhibitor RG1678: results of the proof-of-concept study for the treatment of negative symptoms in schizophrenia. Paper presented at: *49th Annual Meeting of the American College of Neuropsychopharmacology*, Miami Beach, Florida (USA). 2010.
36. Alberati D, Moreau J-L, Lengyel J, Hauser N, Mory R, Borroni E, Pinard E, Knoflach F, Schlotterbeck G, Hainzl D. Glycine reuptake inhibitor RG1678: a pharmacologic characterization of an investigational agent for the treatment of schizophrenia. *Neuropharmacology* 2012;62(2):1152-1161.

Chapter 4

Test case under ‘naturalistic’ conditions in schizophrenia

Evaluation of investigational drug perspectives in comparison with the market leading product [1]

[1] This chapter was published as:

Schoemaker, J., Naber, D., Vrijland, P., Panagides, J., & Emsley, R. (2010). Long-term assessment of asenapine vs. olanzapine in patients with schizophrenia or schizoaffective disorder. *Pharmacopsychiatry 43,* 138-146.

DOI: http://dx.doi.org/ 10.1055/s-0031-1295450.

Erratum: *Pharmacopsychiatry, 2011, 44,* 343.

DOI: http://dx.doi.org/10.1055/s-0030-1248313

ABSTRACT

Introduction: We conducted a double-blind 1-year trial of asenapine in patients with schizophrenia or schizoaffective disorder.

Methods: Patients were randomized to asenapine (5 or 10 mg BID; n=913) or olanzapine (10–20 mg QD; n=312), and monitored regularly.

Results: Trial completion rates were 38% with asenapine and 57% with olanzapine; main reasons for discontinuation were withdrawal of consent (22%, 16%) and insufficient response (25%, 14%); fewer discontinuations were due to adverse events (6%, 7%). Mean weight gain was 0.9 kg with asenapine, and 4.2 kg with olanzapine. Extrapyramidal symptoms reported as adverse events were more common with asenapine. Mean reductions in PANSS total score with asenapine and olanzapine were –21.0 and –27.5 ($P<0.0001$); the exclusion of patients who had previous poor experience with olanzapine may have biased the results in favor of olanzapine. Scores on the Subjective Wellbeing on Neuroleptics scale and functionality measures were similar between groups.

Conclusion: Asenapine was well tolerated over one year of treatment, causing less weight gain than olanzapine but more frequent extrapyramidal symptoms. PANSS total score improved with both agents; the improvement was greater with olanzapine than with asenapine using last observations carried forward but not in an observed-case analysis.

INTRODUCTION

Schizophrenia and schizoaffective disorder are serious mental disorders that impair patients' ability to function. Schizophrenia is characterized principally by positive and negative symptoms,[1] although nearly 70% of patients exhibit cognitive dysfunction and often have depressive symptoms as well.[2,3] Schizoaffective disorder is characterized by a major depressive, manic, or mixed episode in addition to symptoms of schizophrenia.[1]

First-generation (conventional) and second-generation (atypical) antipsychotics are effective in treating the positive symptoms associated with these disorders, but efficacy for negative symptoms and cognitive dysfunction is an unmet need. Considerable accumulated data suggest that antipsychotics differ more in terms of safety and tolerability than efficacy.[4,5]

Asenapine is an atypical antipsychotic approved in the United States for the acute treatment of schizophrenia in adults and for manic or mixed episodes of bipolar I disorder with or without psychotic features in adults. Its human receptor signature is characterized by strong affinity for serotonin receptor subtypes (5-HT_{1A}, 5-HT_{1B}, 5-HT_{2A}, 5-HT_{2B}, 5-HT_{2C}, 5-HT_5, 5-HT_6, 5-HT_7), dopamine receptor subtypes (D_1, D_2, D_3, D_4), α-adrenergic receptors, and histaminic receptors, but no appreciable affinity for muscarinic receptors.[6] Asenapine is administered as a fast-dissolving sublingual tablet for rapid absorption through the oral mucosa.

The efficacy of asenapine for symptoms of acute schizophrenia has been demonstrated in two 6-week trials.[7,8] In both studies, asenapine was significantly better than placebo as measured by improvements on the Positive and Negative Syndrome Scale (PANSS)[9] total score and was generally well tolerated, showing a low incidence of clinically significant weight gain, hyperprolactinemia, and extrapyramidal symptoms (EPS).

The goal of this study was to assess the long-term safety and efficacy of asenapine in a flexible-dose regimen in patients with schizophrenia or schizoaffective disorder, and to explore the safety and efficacy of asenapine in comparison with olanzapine.

METHODS

This randomized, double-dummy, double-blind, flexible-dose, 52-week phase III trial (clinicaltrials.gov registry: NCT00212784) was conducted at 102 sites in Australia, Belgium, the Czech Republic, France, Germany, the Netherlands, Poland, Russia, South Africa, Spain, and the United Kingdom from September 2003 through February 2006. The trial protocol was approved by the independent ethics committee at each site, and all patients provided written informed consent.

Patients

Inpatients or outpatients aged 18 years or older (no upper age limit) were eligible for this trial if they had a Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) diagnosis of schizophrenia or schizoaffective disorder; a PANSS total score ≥60, including scores of ≥4 on at least 2 of 5 items on the PANSS positive subscale at screening and baseline; and a Clinical Global Impression–Severity of Illness (CGI-S)[10] score of ≥4 (moderately ill; score of 5 indicates severe illness) at baseline. Previously treated patients had to have a history of a positive response to an antipsychotic other than clozapine; treatment-naïve patients also were eligible.

Patients with a history of inadequate response to or intolerable adverse effects with olanzapine were excluded, as were those who were using disallowed concomitant medications, such as mood stabilizers or a second antipsychotic, and those responding well to current antipsychotic treatment and who had no clinical need to stop or change treatment. Patients also were excluded if they had a score greater than "mild" on any item of the Abnormal Involuntary Movement Scale[11] or were considered by the investigator to be at risk of self-harm or harm to others.

Patients meeting eligibility criteria were randomly assigned (in a 3:1 ratio) to treatment with asenapine 5 mg twice daily (BID) or olanzapine 10 mg once daily (QD). Asenapine (or matching placebo) was provided in sublingual tablets whereas olanzapine (or matching placebo) was provided in oral capsules. To maintain blinding, a double-dummy design was used; patients in the asenapine group also received once-daily oral placebo and those in the olanzapine group also received twice-daily sublingual placebo. After the first 7 days, dosages were flexible (asenapine 5 or 10 mg BID; olanzapine 10 or 20 mg QD) for the remaining

treatment period. During the study period, concomitant treatment with hypnotics, anxiolytics, anticholinergics and antidepressants (other than tricyclic agents or monoamine oxidase inhibitors) was allowed.

Safety and Tolerability Assessments

Vital signs and body weight were assessed at each visit (screening, baseline, weeks 1, 2, 3, 4, 6, 8, and then every 4 weeks through week 52). Physical examinations were conducted at screening and weeks 6, 24, and 52; electrocardiograms were performed at screening and weeks 3, 6, 24, and 52. Laboratory tests (hematology, biochemistry, urinalysis) were done at screening and weeks 1, 3, 6, 16, 24, 32, 40, and 52.

Tolerability was assessed in terms of adverse events, either reported by the patient or observed by the investigator during patient visits, up to 7 days after the last dose (30 days for serious adverse events). Extrapyramidal symptoms were assessed using the Barnes Akathisia Rating Scale,[12] the Simpson-Angus Scale,[13] and the Abnormal Involuntary Movement Scale[11] at baseline and weeks 1, 3, 6, 16, 24, 32, 40, and 52; EPS-like symptoms were reported as adverse events when considered clinically relevant by the investigator.

Efficacy and Other Assessments

Changes in various symptom domains and overall disease severity for up to 3 days after the last dose were assessed using the PANSS,[14] CGI-Improvement (CGI-I),[10] and CGI-S. CGI assessments were performed at baseline and at every post-baseline assessment to endpoint. PANSS assessments were performed at screening and baseline, and at weeks 2, 4, 6, 8, 12, 20, 28, 36, 44, and 52. Other assessments included the Subjective Well-being Under Neuroleptic Treatment (SWN) scale[15] and the physical and mental component scales of the Medical Outcomes Study 12-Item Short Form (SF-12) health survey.[16] Assessments using the SWN and SF-12 were performed at baseline and at weeks 8, 20, 28, 36, 44, and 52. At the end of treatment, both the patient and the investigator used a 5-point scale, from "much better" to "much worse," to rate their satisfaction with treatment compared with previous antipsychotic medication.

Statistical Analysis

Safety data were obtained from all randomized patients who received ≥1 dose of study medication (all-patients-treated, or safety population). Efficacy was assessed using data from the intent-to-treat (ITT) population, which comprised treated patients who had a baseline PANSS assessment and ≥1 post-baseline PANSS assessment. Olanzapine was used as active control, but no formal power calculations were done to establish non-inferiority or superiority of asenapine versus olanzapine. The exploratory analysis of efficacy utilized an analysis of covariance (ANCOVA, with baseline as a covariate, treatment group as a fixed factor, and site as a random factor) on the change from baseline at each time point in the ITT population (using last observations carried forward [LOCF]) and in observed cases. Discontinuation rates by treatment group were compared using a log-rank test, stratified for site.

RESULTS

Of 1377 patients screened, 1225 were randomly assigned to treatment and 1219 received ≥1 dose of study medication. Baseline demographics and clinical characteristics were similar between the treatment groups (Table 4.1). In the majority of patients in both groups, the diagnosis was a recurrent episode of paranoid schizophrenia; most patients (82%) had no history of suicide attempt. In both groups, more than half the patients were men; 93% of patients were white and 6% were black. In both groups, 61%–62% of patients were single, 20%–21% were married, and 15% were divorced; 43%–44% lived with their parents, 19%–20% lived with their spouses, and 18%–20% lived alone; and 30%–32% were unemployed while 14%–16% earned some form of income.

Patient disposition, including the number of patients completing the trial in each treatment group, is shown in Figure 4.1. A total of 350 (38%) patients on asenapine and 178 (57%) patients on olanzapine completed the trial ($P<0.0001$). The most frequent reasons for premature discontinuation were withdrawal of consent (22% with asenapine, 16% with olanzapine) and worsening of disease/lack of response (25%, 14%). The frequency of discontinuations because of adverse events not related to worsening of disease or inadequate response to treatment was much lower in both treatment groups (6% with asenapine, 7% with olanzapine).

At study end, the final dose in the asenapine group was 5 mg BID for 53% of patients and 10 mg BID for 47% of patients. In the olanzapine group, the final dose was 10 mg QD for 59% of patients and 20 mg QD for 41% of patients. The mean (standard deviation ± SD) daily exposure to each study drug was 13.5±5.1 mg for asenapine and 13.6±4.5 mg for olanzapine.

The most commonly used concomitant medications in both treatment groups were lorazepam (24% in the asenapine group, 21% in the olanzapine group), acetaminophen (10%, 9%), clonazepam (10%, 6%), zolpidem (10%, 6%), diazepam (10%, 8%), phenazepam (8%, 6%), and zopiclone (6%, 5%). At endpoint, a higher percentage of patients in the asenapine group were taking hypnotics, anxiolytics, or anticholinergics compared with the olanzapine group; use of antidepressants was similar between groups (Table 4.2).

Safety and Tolerability

Data on adverse events reported over the course of the year-long trial are summarized in Table 4.3. The incidence of treatment-emergent adverse events was 82% in both groups. The incidence of adverse events that were, in the investigators' judgment, treatment-related was 60% for asenapine and 61% for olanzapine. Most reported adverse events were rated as mild or moderate.

For the asenapine and olanzapine groups, respectively, incidence rates of treatment-emergent serious adverse events were 19% and 12%, and rates of treatment-related serious adverse events were 6% and 2%. The mortality rate was <1% in both groups: there were 7 deaths in the asenapine group (5 by suicide, among 11 attempts) and 1 death in the olanzapine group (by suicide, among 6 attempts). None of the deaths that occurred in the course of the trial was considered treatment-related.

Frequently reported adverse events are summarized in Table 4.4. As shown, most of the treatment-emergent adverse events were considered to be treatment-related, except for worsening psychiatric condition and insomnia; for these adverse events, the overall incidence (treatment-emergent) was substantially higher than the incidence that was considered treatment-related.

Routine laboratory analyses showed no signs of serious metabolic changes, blood dyscrasia, or persistent abnormalities in endocrine, renal, or liver function. Transient abnormalities in liver function tests were seen more frequently with olanzapine. Mean changes from baseline for alanine aminotransferase (ALT), aspartate aminotransferase (AST),

and gamma-glutamyl transpeptidase were smaller with asenapine than with olanzapine (Table 4.5). Markedly abnormal increases (to levels >3 times upper limit of normal) in ALT were seen in 3% of the asenapine group and 11% of the olanzapine group; for AST, the percentages were 1% and 3%, respectively.

Plasma prolactin levels decreased from elevated levels at baseline in both treatment groups (Figure 4.2). No notable changes or between-group differences were seen in measures of total cholesterol or glucose, but triglyceride levels rose substantially with olanzapine and declined slightly with asenapine (Table 4.5). There were no other notable findings regarding hematology, urinalysis, or biochemistry variables.

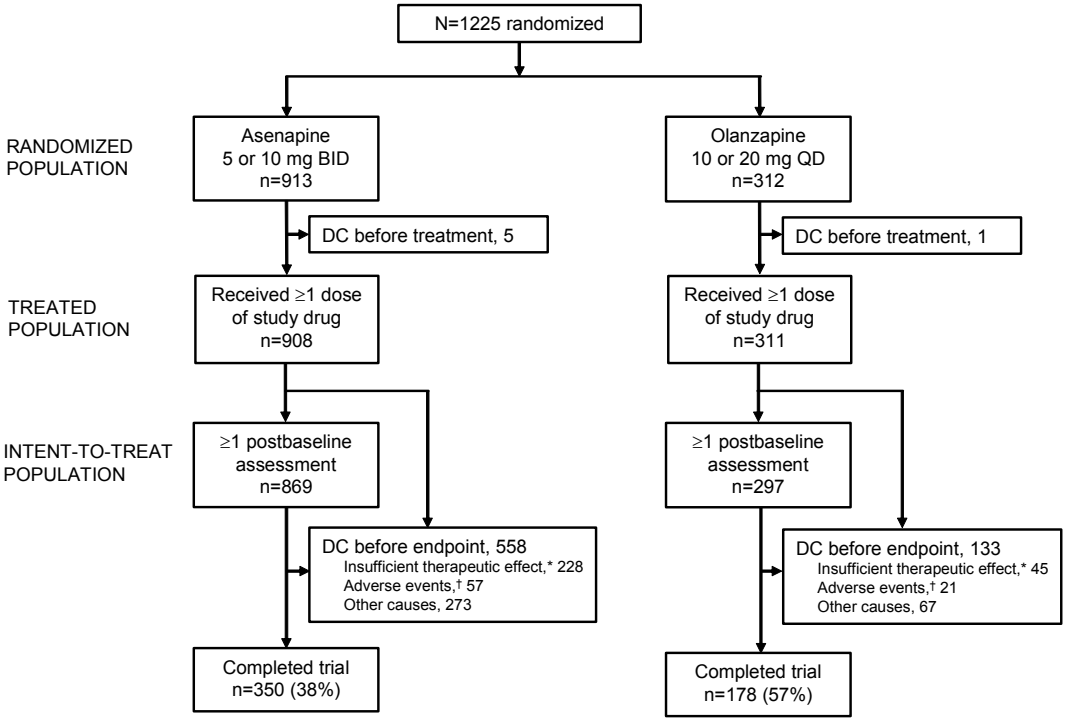


Figure 4.1

Patient disposition.

* Represents combined discontinuations due to lack of efficacy and due to adverse events related to worsening of disease.

† Represents discontinuations due to adverse events not related to worsening of disease

Table 4.1

Demographics and clinical characteristics at baseline, safety population

	Asenapine (n=908)	Olanzapine (n=311)
Men, n (%)	475 (52)	182 (59)
Age, y		
Mean ± SD	36.8±11.8	36.2±12.4
Range	16–71	18–81
Weight, kg, mean ± SD	74.6±15.8	73.6±14.4
Body mass index, kg/m^2, mean ± SD	25.6±4.9	25.2±4.6
DSM-IV-TR diagnosis, n (%)		
Schizophrenia	786 (87)	273 (88)
Paranoid	701 (77)	247 (79)
Disorganized	38 (4)	12 (4)
Catatonic	4 (<1)	1 (<1)
Undifferentiated	43 (5)	13 (4)
Schizoaffective disorder	122 (13)	38 (12)
Any previous schizophrenia episode, n (%)		
Yes	855 (94)	281 (90)
No	53 (6)	30 (10)
Duration of current episode, n (%)		
<2 week	118 (13)	44 (14)
<1 month	235 (26)	64 (21)
1 to <6 months	399 (44)	136 (44)
≥6 to 12 months	83 (9)	43 (14)
>12 months	73 (8)	24 (8)
Hospitalized at baseline, n (%)		
Yes	436 (48)	145 (47)
No	471 (52)	166 (53)
CGI–Severity of illness, n (%)		
Mildly ill	1 (<1)	0 (0)
Moderately ill	312 (34)	113 (36)
Markedly ill	456 (50)	146 (47)
Severely ill	138 (15)	51 (16)
Among the most extremely ill patients	1 (<1)	1 (<1)

PANSS=Positive and Negative Syndrome Scale; SF-12=Medical Outcomes Study 12-Item Short Form; SWN=Subjective Well-being under Neuroleptic Treatment scale.

Table 4.2

Concomitant medication use at baseline and endpoint by drug class, safety population

	Number (%) of Patients	
	Asenapine (n=908)	Olanzapine (n=311)
Hypnotics		
Baseline	119 (13)	36 (12)
Endpoint	137 (15)	33 (11)
Anxiolytics		
Baseline	266 (29)	89 (29)
Endpoint	279 (31)	63 (20)
Antidepressants		
Baseline	8 (1)	5 (2)
Endpoint	114 (13)	39 (13)
Anticholinergics		
Baseline	8 (1)	1 (<1)
Endpoint	57 (6)	5 (2)

Table 4.3

Incidence and severity of adverse events (AEs), safety population

	Number (%) of Patients	
	Asenapine (n=908)	Olanzapine (n=311)
All treatment-emergent AEs	749 (82)	254 (82)
- Mild	178 (20)	67 (22)
- Moderate	406 (45)	135 (43)
- Severe	165 (18)	52 (17)
All treatment-related AEs	548 (60)	190 (61)
- Mild	183 (20)	63 (20)
- Moderate	285 (31)	94 (30)
- Severe	80 (9)	33 (11)
Serious AEs		
- All serious AEs	174 (19)	36 (12)
- Treatment-related serious AEs	50 (6)	7 (2)
Suicides		
- Attempts	11 (1.2)	6 (1.9)
- Completed	5 (<1)	1 (<1)
Deaths		
- All deaths	7 (<1)	1 (<1)
- Treatment-related deaths	0	0
Discontinuations due to AEs		
- Any AE	155 (17)	38 (12)
- Related to worsening symptoms	98 (11)	17 (6)
- Not related to worsening symptoms	57 (6)	21 (7)
- Treatment-related AE	87 (10)	26 (8)
- Serious AE	75 (8)	12 (4)
- Treatment-related serious AE	33 (4)	4 (1)

Table 4.4

Most frequently reported adverse events, safety population

	Number (%)			
	Treatment-emergent		Treatment-related	
	Asenapine (n=908)	Olanzapine (n=311)	Asenapine (n=908)	Olanzapine (n=311)
Weight increase	125 (14)	95 (31)	111 (12)	91 (29)
Worsening schizophrenia	229 (25)	62 (20)	73 (8)	16 (5)
Insomnia	170 (19)	45 (14)	63 (7)	14 (5)
Sedation	81 (9)	33 (11)	75 (8)	32 (10)
Somnolence	89 (10)	30 (10)	83 (9)	30 (10)
Gastrointestinal symptoms	93 (10)	22 (7)	84 (9)	21 (7)
Akathisia	89 (10)	11 (4)	76 (8)	11 (4)

Table 4.5

Mean ± SD change from baseline at endpoint for laboratory variables, safety population

	Asenapine (n=908)		Olanzapine (n=311)	
Hepatic Function	U/L		U/L	
ALT	1.7±24.4		5.4±39.5	
AST	1.4±14.7		2.4±22.6	
γ-GT	–0.7±23.4		5.7±31.8	
Metabolic Function	mg/dL	mmol/L	mg/dL	mmol/L
Total cholesterol	–6.0±35.9	–0.16±0.93	3.2±38.36	0.08±0.99
Triglycerides, fasting	–9.8±92.2	–0.11±1.04	30.4±202.60	0.34±2.30
Glucose, fasting	2.4±21.7	0.13±1.20	3.5±25.14	0.20±1.40

ALT=alanine aminotransferase, also called serum glutamic-pyruvic transaminase (SGPT); AST=aspartate aminotransferase, also called serum glutamic-oxaloacetic transaminase (SGOT); γ-GT=gamma-glutamyl transpeptidase.

Table 4.6

Mean ± SD baseline and change from baseline to endpoint in scores on extrapyramidal symptom rating scales, safety population

	Asenapine (n=908)		Olanzapine (n=311)	
	Baseline	Change*	Baseline	Change*
SAS total score	1.2±2.6	–0.4±2.5	1.2±2.8	–0.7±2.7
BARS total score	0.7±1.6	–0.1±1.9	0.6±1.4	–0.3±1.5
AIMS 7 total score	0.5±1.4	–0.1±1.3	0.3±1.0	–0.2±1.2

AIMS=Abnormal Involuntary Movement Scale; BARS=Barnes Akathisia Rating Scale; SAS=Simpson-Angus Scale.

*Negative change indicates improving symptoms.

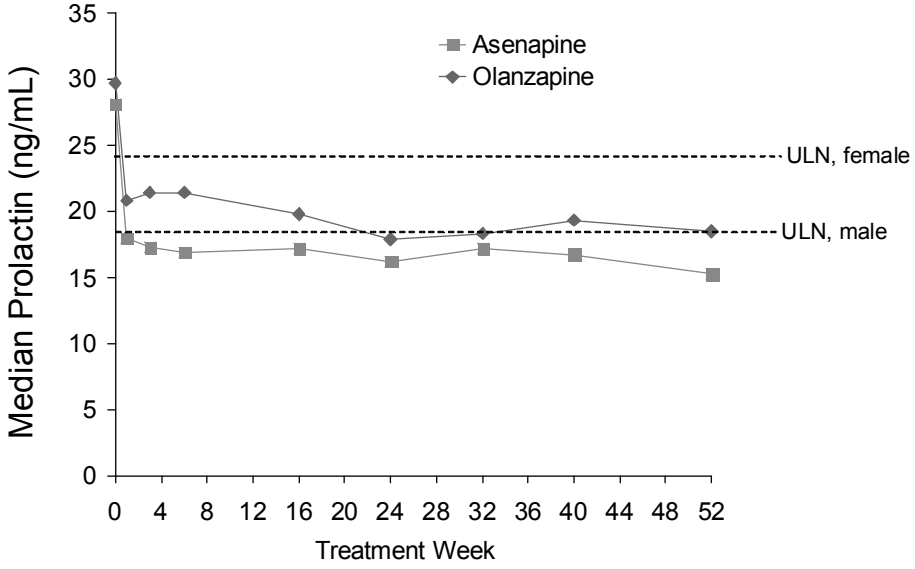


Figure 4.2

Median serum prolactin levels over time in safety population.

ULN=upper limit of normal

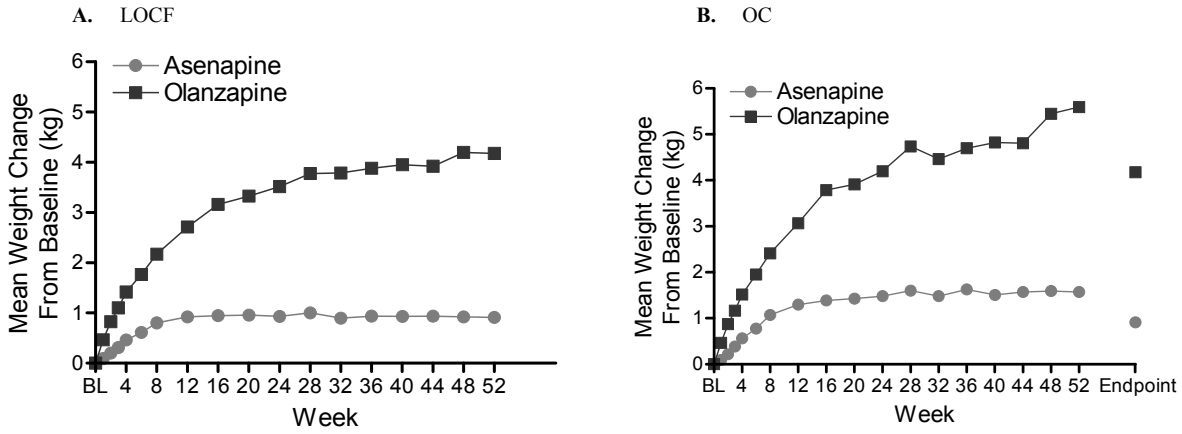


Figure 4.3

Weight change over time in the safety population.

(*a*) Last observations carried forward (LOCF) analysis. (*b*) OC analysis.

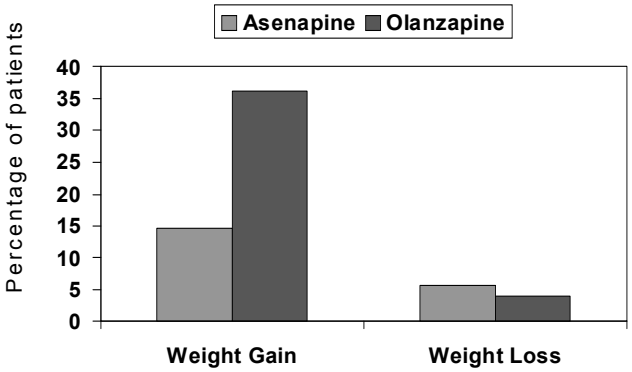


Figure 4.4

Percentage of patients with clinically significant weight gain (≥7% increase from baseline) or weight loss (≥7% decrease from baseline) at endpoint in safety population

Change in weight was greater with olanzapine than asenapine. At baseline, mean ± SD weight was comparable in the asenapine and olanzapine groups: 74.6±15.8 kg (164.5±34.8 lbs) and 73.6±14.4 kg (162.3±31.7 lbs), respectively. At endpoint, using LOCF data from all treated patients, mean ± SD changes were 0.9 ± 4.8 kg (2.0 ± 10.6 Ibs) and 4.2 ± 7.6 kg (9.2 ± 16.8 Ibs), respectively (Figure 3*a*). At week 52, using observed case data, mean ± SD change was l.6 ± 5.7 kg (3.5 ± 12.6 Ibs) for asenapine and 5.6 ± 8.4 kg (12.3 ± 18.5 lbs) for olanzapine (Figure 3*b*). From week 1 onward and at endpoint, between-group differences in change from baseline in body weight were statistically significant, in favor of asenapine ($P<0.01$). Incidence rates of clinically significant weight change at endpoint (≥7% increase or decrease from baseline) are shown in Figure 4.4.

Weight gain was reported as an adverse event in 14% of patients treated with asenapine and in 31% of those treated with olanzapine; most of these cases (12% and 29%) were considered treatment-related. Weight loss was reported as an adverse event in 2% and 3% of patients in the asenapine and olanzapine groups.

In a post hoc subanalyses, weight change was stratified by baseline body mass index (BMI). For BMI <23 kg/m2, mean ± SD changes were 1.72±4.05 kg (3.8±8.9 lbs) with asenapine and 6.98±8.32 kg (15.4±18.3 lbs) with olanzapine, and incidence rates of clinically significant weight gain were 22.0% and 57.1%. For BMI 23–27 kg/m2, changes with asenapine and olanzapine were 0.97±4.37 (2.1±9.6 lbs) and 3.73±5.74 kg (8.2±12.7 lbs), and

rates of significant weight gain were 12.8% and 29.7%. For BMI >27 kg/m2, changes with asenapine and olanzapine were 0.04±5.62 kg (0.1±12.4 lbs) and 1.84±7.94 kg (4.1±17.5 lbs), and rates of significant weight gain were 9.3% and 21.9%.

EPS-like symptoms (movement disorders), usually mild, were reported as an adverse event by 18% of asenapine-treated patients and 8% of olanzapine-treated patients; most of these events (14% and 7%) were considered treatment-related. The most commonly reported type of movement disorder in the asenapine and olanzapine groups was akathisia (treatment-emergent: 10% and 4%; treatment-related: 8% and 4%). EPS, chiefly akathisia, was the reported cause of premature discontinuation in 12 patients taking asenapine (1%) and 3 patients taking olanzapine (1%). Tardive dyskinesia was reported in 3 patients, all taking asenapine (1 of these patients had the condition at baseline). On formal assessment of EPS, mean scores on rating scales decreased from baseline in both groups (Table 4.6). Among observed cases in the asenapine and olanzapine groups, respectively, the proportion of patients using anticholinergic drugs was 7% and 2% at week 6, and 8% and 2% at week 52; and among all treated patients, 6% and 2% at endpoint.

The incidence of abnormalities in vital signs was low in both treatment groups at all time points, and no notable between-group differences were seen. Electrocardiographic abnormalities were reported in 2.4% and 1.3% of patients in the asenapine and olanzapine groups, respectively. Most of these cases were related to lengthening of the baseline rate-corrected QT interval (QTc, Fridericia correction); however, there were no instances of QTc ≥500 msec at any time during treatment.

Efficacy and Other Outcomes

At baseline, patients in the asenapine and olanzapine treatment groups were moderately to severely ill, as reflected by almost identical mean CGI-S scores of 4.8 and PANSS total scores of 92.1 (Table 4.7). Both groups showed improvement over the course of the trial.

In the LOCF analysis in the ITT population, changes (mean ± SD) from baseline in PANSS total score with asenapine and olanzapine were similar at week 6 (–17.9±17.8 and –19.0 ±17.6) but showed a statistically significant difference in favor of olanzapine at endpoint (–21.0±22.8 and –27.5±22.0; P<0.0001). In the observed-case analysis among patients who completed the year-long trial, changes (mean ± SD) in PANSS total score with asenapine and

olanzapine were similar at week 6 (–22.1±15.0 and –21.4±16.0 in 715 and 266 patients, respectively) and also at week 52 (–35.9±16.3 and –35.4±16.2 in 357 and 179 patients).

LOCF and OC assessments of changes in PANSS Marder factor scores and CGI-S scores revealed no significant between-group differences (Table 4.7). On the CGI-I, mean ± SD scores at endpoint using LOCF analysis were 2.9±1.6 and 2.4±1.5, respectively. At endpoint, 48% of patients in the asenapine, and 34% in the olanzapine group had CGI-I scores ≥3 (indicating minimal improvement, no change, or worsening), whereas 52% and 66% of patients had scores <3 (indicating much or very much improvement).

Clinical improvement with treatment was paralleled by increases in mean ± SD SWN total score, from 75.8±17.8 at baseline to 84.4±19.1 at endpoint with asenapine, and from 76.1±16.4 to 85.1±17.7 with olanzapine. No meaningful change from baseline on the mental or physical component of SF-12 was observed at any time during the trial, indicating that patients' overall health status was not notably affected by treatment with either drug.

There were no notable changes within groups or significant between-group differences in living situation, employment, or level of functioning. Among patients enrolled as outpatients, the incidence of hospitalization was higher for asenapine than for olanzapine (6% vs 3%), although the total number of hospital days during the trial was marginally lower for asenapine than for olanzapine (mean ± SD days, 34.9±50.2 vs 36.3±41.6).

At study end, 34% of asenapine-treated patients and 37% of investigators considered asenapine much better than previous antipsychotic medication; and 40% of olanzapine-treated patients and 48% of investigators considered olanzapine much better than previous treatment.

Table 4.7

Mean ± SD baseline and change from baseline on efficacy measures, intent-to-treat population

	Last Observations Carried Forward (LOCF)					Observed Cases (OC)				
	Asenapine (n=869)		Olanzapine (n=297)			Asenapine (n=357)		Olanzapine (n=179)		
	Baseline	Change at Endpoint	Baseline	Change at Endpoint	*P* value*	Baseline	Change at Endpoint	Baseline	Change at Endpoint	*P* value*
PANSS Total score	92.1±13.3	–21.0±22.8	92.1±13.4	–27.5±22.0	<0.001	91.5±13.4	–35.9±16.3	91.0±12.7	–35.4±16.2	0.881
PANSS Marder factors										
Positive symptoms	27.6±4.76	–7.9±7.67	27.6±4.79	–10.0±7.75	<0.001	27.3±4.60	–12.7±5.78	27.2±4.57	–12.6±5.87	0.809
Negative symptoms	22.8±5.74	–4.6±6.54	23.2±5.81	–6.0±6.23	<0.001	22.9±5.60	–7.7±5.75	22.9±5.57	–7.4±5.77	0.467
Disorganized thought	20.9±4.17	–4.4±5.36	21.2±4.28	–5.9±5.29	<0.001	20.9±4.10	–7.3±4.26	20.9±3.87	–7.5±4.27	0.346
Hostility / Excitement	9.4±3.12	–1.5±4.11	9.2±2.99	–2.4±3.72	<0.001	9.3±3.17	–3.4±3.19	9.0±3.04	–3.5±2.94	0.119
Anxiety / Depression	11.2±3.25	–2.7±3.70	11.0±3.39	–3.3±3.79	<0.001	11.0±3.07	–4.6±2.96	11.0±3.08	–4.5±3.13	0.754
CGI–Severity of Illness	4.8±0.68	–1.2±1.35	4.8±0.71	–1.6±1.35	<0.001	4.7±0.67	–2.0±1.10	4.7±0.68	–2.1±1.11	0.260

* Based on analysis of covariance (ANCOVA) with baseline as a covariate, treatment group as a fixed factor, and site as a random factor.

†Baseline in the subset of patients who completed the assessment at week 52.

CGI=Clinical Global Impressions; PANSS=Positive and Negative Syndrome Scale.

DISCUSSION

Adverse events can decrease medication adherence and consequently increase the risk of relapse and costly re-hospitalization, impair patients' health-related quality of life, be stigmatizing (particularly excessive weight gain and abnormal involuntary movements), and increase medical comorbidity and mortality.[17] Thus, one of the goals of this study was to evaluate the long-term safety of asenapine in the treatment of patients with schizophrenia or schizoaffective disorder.

In this year-long trial, asenapine was generally well tolerated. Overall discontinuation rates, however, were higher with asenapine (62%) than with olanzapine (43%), but mostly for reasons other than safety and tolerability. Still higher overall discontinuation rates have been observed in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE), an 18-month study in which 1432 patients with chronic schizophrenia were randomly assigned to double-blind treatment with olanzapine (7.5–30 mg/day), perphenazine (8–32 mg/day), quetiapine (200–800 mg/day), or risperidone (1.5–6.0 mg/day).[18] Overall discontinuation rates in CATIE were lower with olanzapine (64%) than with the other antipsychotics (74%–82%). Discontinuation rates due to lack of efficacy in CATIE were also lower with olanzapine (15%) than with other treatments (24%–28%), and those rates are consistent with the corresponding rates in our study (14% discontinuation rate due to lack of efficacy with olanzapine, 25% with asenapine). Fairly high discontinuation rates are not unusual in schizophrenia trials, and withdrawals because of adverse events are usually less frequent with second-generation antipsychotics than with first-generation agents, at least in flexible-dose trials.[19] Although discontinuations due to insufficient therapeutic effect accounted for 41% and 34% of withdrawals in the asenapine and olanzapine groups, respectively, both drugs improved mean PANSS total scores and more than half of patients in each group had endpoint CGI-I scores indicating that their symptoms were much or very much improved.

Changes on the PANSS Marder factors paralleled the changes on the PANSS total score. On LOCF analysis, changes were numerically larger with olanzapine than with asenapine, whereas the OC analysis showed no notable between-group differences, and the changes in both groups were notably larger than those seen on the LOCF analysis.

The improvement in patients' subjective sense of wellbeing, as reflected in the increased scores on the SWN, is especially important, because the subjective sense of improvement in

the early stages of treatment is associated with a greater likelihood of achieving clinical remission,[20,21] which is likely correlated with improved adherence to treatment.[22] In this study, the improvement on SWN scores with asenapine was comparable to the improvement with olanzapine. In some previous studies, improvements on SWN scores were larger with olanzapine than with other atypical antipsychotics,[15,23] although a large double-blind study found no difference in improvement in SWN scores between olanzapine and risperidone.[24] As is often the case when clinical trials produce conflicting outcomes, the differences may be due to variability in patient populations, medication dosages, and other factors. It is also interesting to note that the improvements seen in SWN scores in the present study were not matched by any notable change in SF-12 scores. Although both instruments reflect the patient's state of wellbeing, the SF-12 is a generic scale whereas SWN was specifically developed for schizophrenia or psychosis patients being treated with antipsychotics. Thus, the SWN would be expected to have greater sensitivity in detecting treatment effects in this patient population.

The concomitant use of anxiolytics and sedative-hypnotics was higher in the asenapine group than in the olanzapine group, which may in part reflect a greater sedative effect with olanzapine or a higher incidence of worsening symptoms reported as an AE with asenapine, or both. The use of anticholinergic drugs was also higher with asenapine (see below), but the use of antidepressants was similar with both treatments.

In our study, mean weight gain at study end was greater in the olanzapine group, as was the percentage of patients with clinically significant weight gain. These findings applied not only to the study population as a whole but also to every BMI category when weight change was stratified by baseline BMI. For both treatment groups, weight gain was greatest among patients with the lowest BMI at baseline. Change in weight over time in observed cases (Figure 4.3) suggests that patients taking olanzapine were still gaining weight at study end, while weight gain had plateaued in patients treated with asenapine. Thus, long-term weight gain appears to be lower with asenapine than with olanzapine.

We also noted differential effects on liver enzymes; increases in ALT and AST were greater with olanzapine than with asenapine over the 52-week study period. Transient increases in hepatic enzymes with olanzapine have previously been reported, with clinically significant increases in ALT documented in about 2% of patients taking olanzapine.[25] On other laboratory measures, we found that total cholesterol and triglyceride levels increased

with olanzapine but decreased with asenapine; the increase in triglycerides in patients treated with olanzapine was noteworthy. Olanzapine has long been associated with adverse metabolic effects,[17] and a recent comprehensive review has concluded that these effects show correlation with serum drug levels, implying a dose relationship.[26] Small increases in fasting glucose were seen in both treatment groups.

A decreased risk of EPS at therapeutically effective doses is a defining feature of second-generation antipsychotics compared with first-generation agents.[25] In this regard, asenapine appears similar to current second-generation antipsychotics. The incidence of akathisia (the most frequently reported form of EPS) among patients treated with asenapine in our study was higher than the incidence with olanzapine (10% vs 4%), which may in part account for the more frequent concomitant use of anticholinergic drugs in the asenapine group. The incidence of akathisia among patients treated with first-generation antipsychotics is substantially higher (approximately 50%),[25] but there are no data directly comparing akathisia incidence with asenapine versus first-generation antipsychotics. Tardive dyskinesia is a more serious condition that can develop with long-term antipsychotic treatment and can persist even after medication is withdrawn. However, the incidence of tardive dyskinesia appears to be relatively low with second-generation antipsychotics.[25] These drugs conferred a lower risk for tardive dyskinesia (0.9%) than first-generation antipsychotics (3.8%) at 6 months treatment in patients from the European Schizophrenia Outpatient Health Outcomes (SOHO) study,[27] which is consistent with our experience with asenapine in this trial.

Strengths of our study include the large sample size, double-blind design, and long duration of treatment. One weakness of our study is the double-dummy design, requiring patients (who tend to show poor compliance) to take twice as many doses as otherwise would be necessary. Also, the use or non-use of different antipsychotics before enrollment and/or use of concomitant medication could have affected the outcomes.

Another limitation, common to all randomized controlled trials, is that participants may not be representative of "real world" patients with schizophrenia. For example, the exclusion of patients deemed to be at risk of self-harm or harm to others may limit the applicability of these results to the real-world population of patients with schizophrenia, not only in terms of therapeutic response but also in terms of treatment adherence, which may be compromised by hostile and uncooperative interactions with professional caregivers.[28] The low mortality rate observed in this trial (<1% with both treatments) may also reflect the exclusion of individuals

considered at-risk for suicide. Finally, the exclusion of patients who had not responded well or could not tolerate previous treatment with olanzapine may have produced a bias in favor of olanzapine.

To summarize, assessment of safety and tolerability in this year-long trial indicated that EPS (mainly akathisia) were more frequent with asenapine, but weight gain was more common and more pronounced with olanzapine. Improvement on the primary measure of clinical efficacy, PANSS total score, was greater with olanzapine than with asenapine in an LOCF analysis but not in an observed-case analysis. There were no notable between-group differences in measures of subjective wellbeing or functional status.

ACKNOWLEDGMENTS

This work was supported by Schering-Plough. Editorial support was provided by Complete Healthcare Communications, Inc., and funded by Schering-Plough.

The authors thank the following investigators who participated in this trial: Australia: Rajan Thomas (Box Hill); Nicholas Keks (Ringwood); Michael Theodoros (New Farm); Assen Jablensky (Perth); Kevin McNamara (Southport); Cherrie Galletly (Adelaide); Brett Emmerson (Fortitude Valley); Keith Muir (Nambour); Kate Redman (Maroochydore); Belgium: Claudine Mertens (Sint Denijs Westrem); Serge Seghers (Kortrijk); Jef Hulselmans (Antwerpen); Eugeen De Bleeker (Sint Niklaas); Geert De Bruecker (Lede); Antonio Gazziano (Tielt); Hans Hellebuyck (Ieper); Erik Buntinx (Alken); Joseph Peuskens (Kortenberg); Lieven De Weirdt (Sint Niklaas); Czech Republic: Jaroslave Boucek (Olomouc); Marius Byss (Havlíèkùv Brod); Eva Ceskova (Brno Bohunice); Zdenka Stanková (Ústí nad Labem); Pavel Mohr (Praha); Vlasta Hanušková (Opava); Vlastimil Tichy (Praha); Eva Soukupova (Plzen); Jiri Bilik (Olomouc); David Holub (Mìlník); Zdenek Solle (Praha); Hana Hornochova (Brno); Jirí Krombholz (Praha); Iva Klanova (Sadská); France: Daniel Dassa (Marseille); Jean Tignol (Bordeaux); Mocrane Abbar (Nîmes); Jean Dalery (Bron); Georges-Pierre Noel Badet (Dijon); Naima Kienlen (Rouffach); Frédéric Kidichian (Rouffach); Daniel Bonnaffoux (Dole); Michel Patris (Strasbourg); Georges Zaykine (Narbonne); Joel Thomas (Saint Nazaire); Germany: Wolfgang Maier (Bonn); Claus

Normann (Freiburg); Isabella Heuser (Berlin); Werner Felber (Dresden); Frank Matakas (Koln); W. Kissling (München); Harald-J. Freyberger (Stralsund); Hans-Juergen Moeller (München); Heinrich Sauer (Jena); Thomas Barth (Chemnitz); Peter Bräunig (Chemnitz); The Netherlands: C. Doorschot (Vught); Hans Hovens (Poortugaal); Poland: Jerzy Matysiakiewicz (Tarnowskie Gory); Mieczyslav Janiszewski (Torun); Maciej Czerwinski (Torun); Aleksander Araszkiewicz (Bydgoszcz); Ryszard Wardenski (Zuromin); Andrzej Rajewski (Poznan); Jaroslaw Strzelec (Tuszyn); Zbigniew Skurczyñski (Zabki); Irena Krupka-Matuszczyk (Katowice); Andrzej Kokoszka (Warszawa); Michal Kujawski (Zabrze); Janusz Janczewski (Chelmno); Anna Grzywa (Lublin); Ireneusz Kaczorowski (Belchatów); Przemyslaw Bogacki (Skorzewo); Russian Federation: Boris Andreyev (Leningrad); Sergey Elkin (St. Petersburg); Mikhail Popov (St. Petersburg); Mikhail Ivanov (St. Petersburg); Yury Popov (St. Petersburg); Evgeny Snedkov (St. Petersburg); Vladimir Tochilov (St. Petersburg); Vladimir Vilyanov (Saratov); Vladislav Shamrey (St. Petersburg); Yury Aleksandrovsky (Moscow); Nikolay Neznanov (St. Petersburg); Margarita Morozova (Moscow); Alexander Okhapkin (Smolensk); Mikhail Sheyfer (Samara); Kausar Yakhin (Kazan); South Africa: George Hart (Johannesburg); Don Wilson (Cape Town); Hilda Russouw (Western Cape); Robert Bothwell (Cape Town); Paul Strong (Cape Town); C. Weyers Westdene (Bloemfontein); Carrllo Gagiano (Westdene, Bloemfontein); P.J. Pretorius (Pretoria West); Merryll Vorster (Krugersdorp); Merryll Vorster (Johannesburg); Spain: Alejandro Pons (Barcelona); Jesús de la Gándara (Burgos); Manuel Angel Franco (Zamora); Miguel Angel Gonzalez Torres (Cuenca); Rosario de Arce (Madrid); United Kingdom: Craig Ritchie (London); Ann Mortimer (Lincolnshire).

REFERENCES

1. Association AP. Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision). 2000.
2. Kurtz MM. Neurocognitive impairment across the lifespan in schizophrenia: an update. *Schizophrenia Research* 2005;74(1):15-26.
3. Möller H-J. Occurrence and treatment of depressive comorbidity/cosyndromality in schizophrenic psychoses: conceptual and treatment issues. *World Journal of Biological Psychiatry* 2005;6(4):247-263.
4. Motlová L, Španiel F, Höschl C, Balon R. Are there any differences in the efficacy among second generation antipsychotics in the treatment of schizophrenia and related disorders? *Annals of Clinical Psychiatry* 2007;19(2):133-143.
5. Tandon R, Fleischhacker WW. Comparative efficacy of antipsychotics in the treatment of schizophrenia: a critical assessment. *Schizophrenia Research* 2005;79(2):145-155.
6. Shahid M, Walker GB, Zorn SH, Wong EH. Asenapine: a novel psychopharmacologicagent with a unique human receptor signature. *Journal of Psychopharmacology* 2008;23(1):65-73.
7. Potkin SG, Cohen M, Panagides J. Efficacy and tolerability of asenapine in acute schizophrenia: a placebo-and risperidone-controlled trial. *Journal of Clinical Psychiatry* 2007;68(10):1,478-1500.
8. Kane J, Zhao J, Panagides J. Efficacy and safety of asenapine in patients with acute exacerbation of schizophrenia. Paper presented at: American Psychiatric Association; May 3-8, 2008; Washington, DC.
9. Kay SR, Flszbein A, Opfer LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 1987;13(2):261-276.
10. Guy W. Clinical global impression scale. *The ECDEU Assessment Manual for Psychopharmacology-Revised Volume DHEW Publ No ADM* 1976;76(338):218-222.
11. Guy W. Abnormal Involuntary Movement Scale. *The ECDEU Assessment Manual for Psychopharmacology-Revised Volume DHEW Publ No ADM* 1976;76(338):534-537.
12. Barnes TR. A rating scale for drug-induced akathisia. *British Journal of Psychiatry* 1989;154(5):672-676.
13. Simpson G, Angus J. A rating scale for extrapyramidal side effects. *Acta Psychiatrica Scandinavica* 1970;45(S212):11-19.
14. Kay SR. Positive-negative symptom assessment in schizophrenia: psychometric issues and scale comparison. *Psychiatric Quarterly* 1990;61(3):163-178.
15. Naber D, Moritz S, Lambert M, Rajonk F, Holzbach R, Mass R, Andresen B, Frank P, Rüdiger H, Reinhard M. Improvement of schizophrenic patients' subjective well-being under atypical antipsychotic drugs. *Schizophrenia Research* 2001;50(1):79-88.
16. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care* 1992:473-483.
17. Haddad PM, Sharma SG. Adverse effects of atypical antipsychotics. *CNS Drugs* 2007;21(11):911-936.
18. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RS, Davis SM, Davis CE, Lebowitz BD. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine* 2005;353(12):1209-1223.
19. Martin JLR, Pérez V, Sacristán M, Rodríguez-Artalejo F, Martínez C, Álvarez E. Meta-analysis of drop-out rates in randomised clinical trials, comparing typical and atypical antipsychotics in the treatment of schizophrenia. *European Psychiatry* 2006;21(1):11-20.
20. De Haan L, van Nimwegen L, van Amelsvoort T, Dingemans P, Linszen D. Improvement of subjective well-being and enduring symptomatic remission, a 5-year follow-up of first episode schizophrenia. *Pharmacopsychiatry* 2008;41:125-128.
21. Lambert M, Schimmelmann BG, Naber D, Schacht A, Karow A, Wagner T, Czekalla J. Prediction of remission as a combination of symptomatic and functional remission and adequate subjective well-being in 2960 patients with schizophrenia. *Journal of Clinical Psychiatry* 2006;67(11):1690-1697.
22. Karow A, Czekalla J, Dittmann RW, Schacht A, Wagner T, Lambert M, Schimmelmann BG, Naber D. Association of subjective well-being, symptoms, and side effects with compliance after 12 months of treatment in schizophrenia. *Journal of Clinical Psychiatry* 2007;68(1):75-80.
23. Naber D, Riedel M, Klimke A, Vorbach EU, Lambert M, Kühn KU, Bender S, Bandelow B, Lemmer W, Moritz S. Randomized double blind comparison of olanzapine vs. clozapine on subjective well-being and clinical outcome in patients with schizophrenia. *Acta Psychiatrica Scandinavica* 2005;111(2):106-115.

24. van Nimwegen LJ, van Beveren NJ, van der Helm MSc M. Effect of olanzapine and risperidone on subjective well-being and craving for cannabis in patients with schizophrenia or related disorders: a double-blind randomized controlled trial. *Canadian Journal of Psychiatry* 2008;53(6):400-405.
25. Conley RR, Kelly DL. Second-generation antipsychotics for schizophrenia: a review of clinical pharmacology and medication-associated side effects. *Israel Journal of Psychiatry and Related Sciences* 2005;42(1):51-60.
26. Simon V, De Hert M. Are weight gain and metabolic side effects of atypical antipsychotics dose dependent? A literature review. *Journal of Clinical Psychiatry* 2009;70(7):1,478-1050.
27. Tenback DE, van Harten PN, Slooff CJ, Belger MA, van Os J. Effects of antipsychotic treatment on tardive dyskinesia: a 6-month evaluation of patients from the European Schizophrenia Outpatient Health Outcomes (SOHO) Study. *Journal of Clinical Psychiatry* 2005;66(9):1130-1133.
28. De Haan L, Van Amelsvoort T, Dingemans P, Linszen D. Risk factors for medication non-adherence in patients with first episode schizophrenia and related disorders; a prospective five year follow-up. *Pharmacopsychiatry* 2007;40(6):264-268.

Chapter 5

Humanitarian extension of experimental treatment in schizophrenia

Evaluation of long-term efficacy & safety under controlled conditions[1]

[1] This chapter was published as:

Schoemaker, J., Stet, L., Vrijland, P., Naber, D., Panagides, J., & Emsley, R. (2012). Long-term efficacy and safety of asenapine or olanzapine in patients with schizophrenia or schizoaffective disorder: an extension study. *Pharmacopsychiatry, 45(5),*196-203. DOI: http://dx.doi.org/ 10.1055/s-0031-1301310.

ABSTRACT

Introduction: Safety and efficacy results, collected in schizophrenia and schizoaffective disorder patients treated up to nearly 3 years, are presented for asenapine and olanzapine.

Methods: Patients completing a 52-week randomized double-blind core study on flexible-dose asenapine (5 or 10 mg BID) or olanzapine (10 or 20 mg QD) could continue treatment until study blind was broken.

Results: 290 patients on asenapine and 150 on olanzapine continued treatment for variable lengths of time (mean ± SD (range) 311.0±146.1 (10−653) d and 327.4±139.6 (15−631) d, respectively). Adverse event (AE) incidence was lower during extension (asenapine, 62%; olanzapine, 55%) than during core study (78%, 80%). In both groups, body weight increase and incidence of extrapyramidal AEs were negligible during the extension. Mean PANSS total score changes during first year of treatment were –37.0 for asenapine and –35.3 for olanzapine, with further change of 1.6 for asenapine and –0.8 for olanzapine at the extension study endpoint..

Conclusions: Clinical stability on asenapine as well as olanzapine was maintained, with few recurrent or newly emerging AEs beyond 1 year of treatment.

INTRODUCTION

Schizophrenia affects approximately 4 to 7 of every 1000 persons, with the median lifetime prevalence per 1000 persons being nearly equal for men and women (3.7 for men, 3.8 for women).[1] Lifetime prevalence of schizoaffective disorder in the general population has been estimated at 0.32%.[2] While psychosocial and support therapies are valuable in helping patients with schizophrenia or schizoaffective disorder function in society, antipsychotic medication is the mainstay of treatment. Antipsychotic medications such as chlorpromazine have been available since the 1950s, but these first-generation antipsychotics (FGAs) are associated with a risk of adverse events (AEs), particularly extrapyramidal symptoms (EPS).[3] The AEs associated with FGAs frequently contribute to poor treatment compliance and consequently poor functional outcome and decreased quality of life. The introduction of second-generation antipsychotics in the 1990s was considered a treatment advance because these agents seemed to be associated to a lesser extent with EPS and prolactin increase, while showing effect sizes of similar magnitude in comparison with FGAs.[3-6]

Asenapine is an antipsychotic with a unique pharmacological profile approved in the United States for treatment of schizophrenia and acute treatment of manic or mixed episodes associated with bipolar I disorder, either as monotherapy or adjunctive therapy with lithium or valproate;[7] asenapine is indicated in the European Union for the treatment of moderate to severe manic episodes associated with bipolar I disorder in adults.[8] In two 6-week trials, asenapine 5 mg twice daily (BID) was superior to placebo in treating acute schizophrenia, as demonstrated by significantly greater decreases from baseline in Positive and Negative Syndrome Scale (PANSS) total score.[9,10]

To evaluate the long-term safety and efficacy of asenapine in combination with functional outcome, a double-blind phase III trial (25517; clinicaltrials.gov registry NCT00212784) was conducted.[11] In this trial, patients with schizophrenia or schizoaffective disorder were randomly assigned with a 3:1 ratio to treatment with asenapine 5 or 10 mg BID or olanzapine 10 or 20 mg once daily (QD) for up to 52 weeks. Overall, asenapine was well tolerated over one year of treatment, causing no clinically relevant metabolic changes and less weight increase than olanzapine (0.9 kg with asenapine vs 5.6 kg with olanzapine). However, asenapine treatment was associated with higher incidence of EPS (18% vs 8% with olanzapine). The most frequently reported treatment-emergent AEs were weight gain

(asenapine, 14%; olanzapine, 31%), worsening psychotic symptoms (25%, 20%), and insomnia (19%, 14%). In both groups, PANSS total score improved from baseline to endpoint. The improvement was significantly greater for olanzapine compared with asenapine with last observation carried forward (LOCF) analysis (mean reductions in the intent-to-treat [ITT] population: –21.0 with asenapine (n=869) vs –27.5 with olanzapine (n=297); $P<0.0001$) but were comparable with observed case (OC) analysis (–35.9 with asenapine vs –35.4 with olanzapine). Upon completion of the 1-year core study, patients were given the opportunity to continue double-blind treatment until the study blind was broken. The objective of this paper is to describe the long-term effects of asenapine in comparison with olanzapine during this extension period and to evaluate the level of change in functional outcome measures in association with clinical response from treatment start to treatment discontinuation.

METHODS

Study Design

This double-blind phase III extension study (25520; clinicaltrials.gov registry NCT00212771) was conducted at a total of 75 sites in Australia, Belgium, Czech Republic, France, Germany, Poland, Russia, South Africa, and Spain between October 2004 and October 2006. The study was conducted in accordance with principles of Good Clinical Practice and was approved by the appropriate institutional review boards and regulatory agencies. All patients signed a written informed consent document or, if illiterate, provided verbal witnessed consent before enrollment in the extension. The extension used a "humanitarian" protocol, allowing patients to continue treatment if they had benefited after 52 weeks of treatment in the core study.

Patients

Inclusion criteria for the core study were age ≥18 years; *Diagnostic and Statistical Manual of Mental Disorders (4th Edition, Text Revision)* diagnosis of schizophrenia or schizoaffective disorder; PANSS total score ≥60, with scores of ≥4 on at least two of 5 items on the PANSS positive subscale at screening and baseline; and a score of ≥4 on the Clinical Global Impression–Severity of Illness scale (CGI-S) at baseline.[11] Patients who completed the core study, had benefitted from treatment in the opinion of investigator and/or patient, and who

wished to continue on double-blind treatment, were eligible for the extension. Patients were excluded from the extension if, in the judgment of the investigator, they had a medical condition or required concomitant treatment that would confound trial results or increase the risk of treatment failure or unacceptable AEs.

Treatment

Patients in the core study were randomly assigned to double-blind treatment with sublingual asenapine 5 mg BID or oral olanzapine 10 mg QD for the first seven days of treatment. A double-dummy design was used to maintain blinded conditions, with patients in the asenapine group also receiving once-daily oral placebo and those in the olanzapine group also receiving twice-daily sublingual placebo. Compliance was assessed by means of tablet count and estimated by the investigators in percentage of prescribed amount taken. After the first seven days, dosages were flexible (asenapine 5 or 10 mg BID; olanzapine 10 or 20 mg QD) for the remaining treatment period. Patients entering the extension continued on the same treatment regimen, with no re-randomization. The extension treatment was truncated in all ongoing patients when results from the core study were unblinded and reported (i.e., 23 months after the first patient had been enrolled in the extension).

No concomitant medications were prohibited during the extension; however, patients were instructed not to take any medications without consulting the investigator. To avoid potential interactions with asenapine, possible cytochrome (CYP) 2D6 inhibitors and drugs metabolized by the CYP2D6 pathway were to be used with caution.

Safety Assessments

Safety assessments included physical examinations and monitoring of changes in vital signs, weight, laboratory variables (e.g., hematology, blood chemistries and enzyme levels, urinalysis, prolactin), and electrocardiographic findings. Tolerability was assessed through AE reports, which were elicited at each visit through questioning and examination of the patient. Extrapyramidal symptoms were reported as AEs, and their occurrence and severity were rated using the Barnes Akathisia Rating Scale, Simpson-Angus Scale, and Abnormal Involuntary Movement Scale.

Efficacy Assessments

The primary efficacy endpoint was change in PANSS total score from core study baseline to extension endpoint. Secondary efficacy assessments included change in scores on PANSS subscales (positive, negative, general psychopathology), PANSS Marder factors (positive, negative, disorganized thought, hostility/excitement, anxiety/depression), CGI-S, Clinical Global Impression–Improvement scale (CGI-I), and Calgary Depression Scale for Schizophrenia (CDSS). Health-related quality-of-life measures included Subjective Well-Being Under Neuroleptics (SWN), Medical Outcomes Study 12-item Short Form (SF-12), and an abbreviated, 4-item version of the Level of Function (LOF) scale, measuring quantity and quality of both social contacts (scored on a scale from 0–3) and useful work (scored on a scale from 0–2).[12,13]

Data Analyses

Summary statistics are provided for demographic and clinical characteristics at baseline of the core study for all patients who were started on extension treatment.

Safety and tolerability assessments are reported for all patients who completed the core study and treated in the extension. The in-treatment period was defined as the first day of drug exposure in the core study until seven days after the last day trial medication was taken. In patients enrolled in the extension, safety/tolerability data are presented showing the change at core study endpoint and the change over the entire treatment period using the baseline of the core study as a reference point.

Efficacy is only descriptively explored in the current study. There was no hypothesis testing for treatment effects because this humanitarian extension study did not re-randomize patients after completion of the core study. The primary aim of study was to assess safety. Efficacy is reported in three descriptive ways, one using the LOCF method to impute for missing data in the extension ITT population (ITT_{ext} = all patients started on extension treatment who had ≥1 PANSS assessment in the extension period), one using LOCF for all patients started on treatment in the core study (ITT_{core} = all patients started on treatment who had ≥1 PANSS assessment post-baseline), and one looking at patients who remained on treatment (OC). Efficacy data are presented for the entire treatment period using baseline and endpoint of the core study as reference points. Descriptive statistics were used for all assessments; formal statistical testing was not performed.

The status of SWN, LOF, and SF-12 as independent measurements of quality of life was explored through calculation of Spearman rank correlations between these parameters. Likewise, the extent to which clinical improvement, as assessed by total PANSS scores, was associated with these functional outcome measurements was determined through calculation of Spearman rank correlations.

RESULTS

Patient Demographics

No significant differences in baseline demographics or clinical characteristics were noted between the groups for patients who were started on extension treatment (Table 5.1).

Patient Disposition

Of the 1225 patients randomized to treatment in the core study (913 to asenapine, and 312 to olanzapine), 528 completed 52 weeks of treatment (Figure 5.1). Only patients who completed the core study, had benefitted from treatment in the opinion of the investigator and/or patient, and who wished to continue on double-blind treatment were eligible for the extension; 440 patients (asenapine, 290; olanzapine, 150) were considered eligible and continued on double-blind treatment during the extension and 326 (asenapine, 203; olanzapine, 123) were still on double-blind treatment when the extension was truncated (Figure 5.1, Table 5.2). The ITT_{ext} population consisted of 414 patients (asenapine, 267; olanzapine, 147). During the extension, 87 patients (30.0%) in the asenapine group and 27 (18.0%) in the olanzapine group discontinued treatment, withdrawal of consent being the most common reason for discontinuation in both groups (Figure 5.2).

Drug Exposure and Patient Compliance

Patients continued treatment for a variable duration (asenapine, 10−653 d; olanzapine, 15−631 d). The mean ± SD duration of extension treatment was 311.0±146.1 days with asenapine and 327.4±139.6 days with olanzapine, and the mean ± SD duration of total exposure (core study + extension) was 676.3±148.1 days with asenapine and 692.5±140.7

days with olanzapine. Mean ± SD daily dosage was 13.4±4.6 mg for asenapine and 13.4±4.6 mg for olanzapine in the extension only and 13.4±4.05 mg and 13.4±4.09 mg for asenapine and olanzapine, respectively, during the entire treatment period. Approximately two thirds of the patients (asenapine, 64.1%; olanzapine, 66.0%) were maintained at a modal dose of 10 mg/d; the remainder received a modal dose of 20 mg/d for each treatment. Treatment compliance (%) was defined as 100 times the number of sublingual tablets and/or capsules not returned by the subject divided by the number of sublingual tablets and/or capsules that the subject should have taken. All tablets and capsules not returned were assumed to be taken. Compliance with treatment was satisfactory (≥75%) in all but 2 patients in each treatment group.

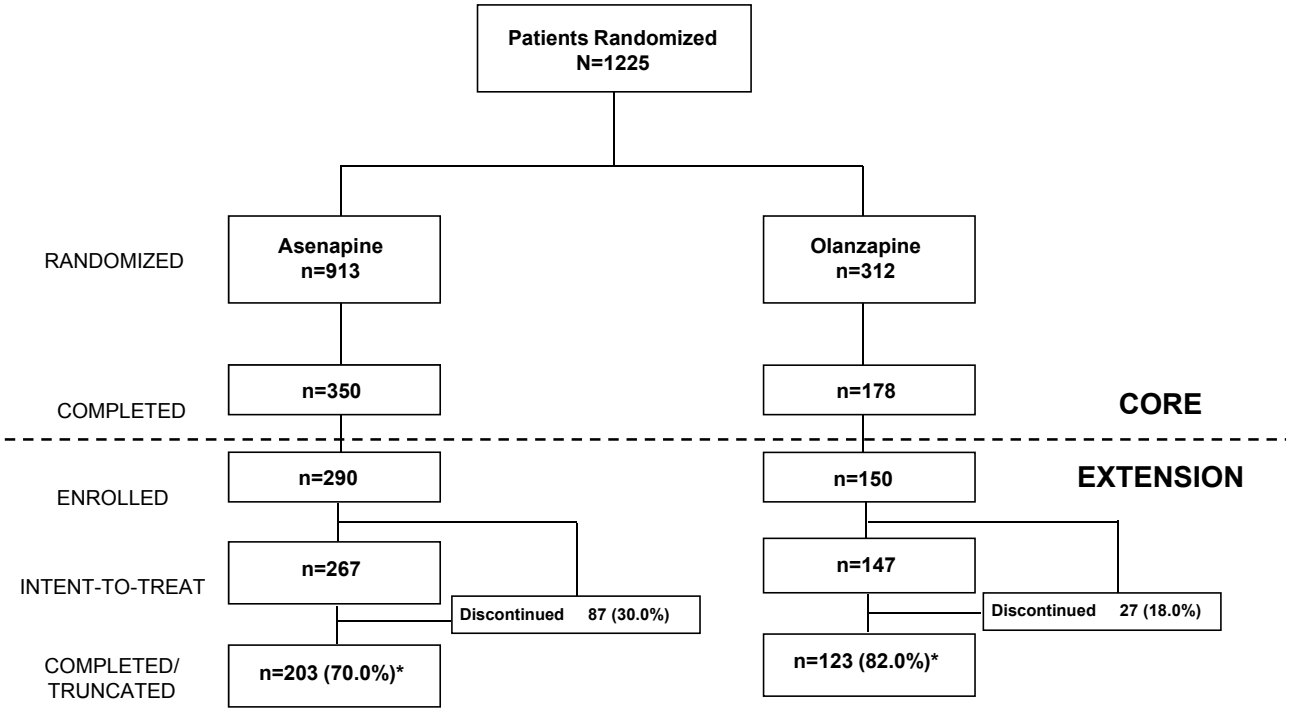


Figure 5.1

Patient disposition. *Designated as completers (i.e., the number of patients in the trial at the time the decision was made to end the extension); completion rate based on enrolled population

Table 5.1

Demographic and clinical characteristics at baseline of the core study for patients started on extension treatment

	Asenapine (n=290)	Olanzapine (n=150)	Total (n=440)
Male sex, n (%)	156 (53.8)	88 (58.7)	244 (55.5)
Age, years			
Mean ± SD	37.1±11.9	36.5±12.6	36.9±12.1
Range	16–68	18–77	16–77
Weight, kg*			
Mean ± SD	74.7±15.4	73.4±15.4	74.2±15.4
Height, cm			
Mean ± SD	171.4±8.8	170.7±9.9	171.1±9.2
Body mass index, kg/m^2*			
Mean ± SD	25.4±4.8	25.1±4.6	25.3±4.7
DSM-IV diagnosis, n (%)			
Schizophrenia			
Paranoid	230 (79.3)	120 (80.0)	350 (79.5)
Disorganized	10 (3.4)	4 (2.7)	14 (3.2)
Catatonic	1 (0.3)	1 (0.7)	2 (0.5)
Undifferentiated	13 (4.5)	7 (4.7)	20 (4.5)
Schizoaffective disorder	36 (12.4)	18 (12.0)	54 (12.3)
Previous schizophrenia episode, n (%)			
Yes	275 (94.8)	140 (93.3)	415 (94.3)
No	15 (5.2)	10 (6.7)	25 (5.7)
Hospitalized at baseline, n (%)			
Yes	108 (37.2)	54 (36.0)	162 (36.8)
No	182 (62.8)	96 (64.0)	278 (63.2)
PANSS total score			
Mean ± SD	90.9±13.1	90.4±12.7	90.7±12.9
CGI-S, n (%)			
Mildly ill	1 (0.3)	0 (0)	1 (0.2)
Moderately ill	125 (43.1)	67 (44.7)	192 (43.6)
Markedly ill	130 (44.8)	66 (44.0)	196 (44.5)
Severely ill	34 (11.7)	17 (11.3)	51 (11.6)

CGI-S=Clinical Global Impression–Severity of Illness scale; DSM-IV=*Diagnostic and Statistical Manual of Mental Disorders (4th Edition, Text Revision)*. *Asenapine, n=289

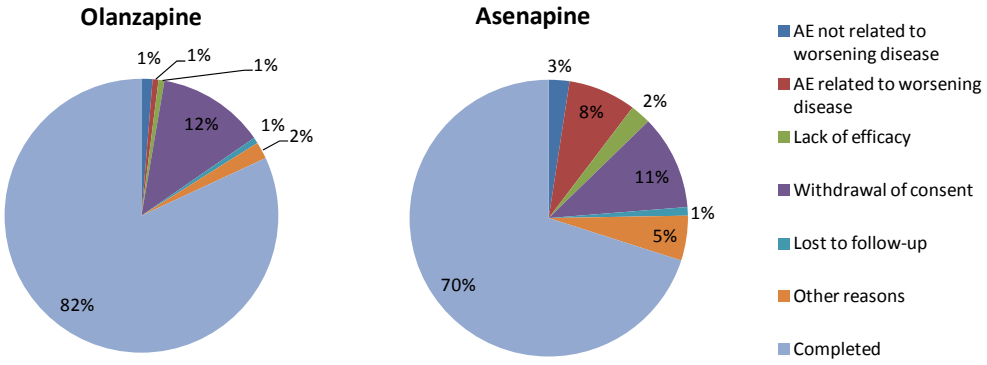


Figure 5.2

Percentages of enrolled population discontinued by reason; AE=adverse event

Table 5.2.

Summary of duration of treatment, all-subjects-started-extension group

Duration of treatment	Asenapine (n=290)	Olanzapine (n=150)
≥365 days	290 (100.0)	150 (100.0)
≥449 days	269 (92.8)	147 (98.0)
≥533 days	245 (84.5)	137 (91.3)
≥617 days	191 (65.9)	101 (67.3)
≥701 days	108 (37.2)	56 (37.3)
≥785 days	70 (24.1)	42 (28.0)
≥869 days	37 (12.8)	25 (16.7)
≥953 days	10 (3.4)	3 (2.0)
≥1037 days	3 (1.0)	0 (0.0)

Concomitant Medication Use

Concomitant psychotropic medication use is presented in Table 5.3. The percentage of patients using antidepressants and anticholinergics increased during the core study for both groups but was fairly stable during the extension. The percentage of patients using concomitant hypnotics during the extension increased for asenapine and remained stable for olanzapine. The percentage of patients using anxiolytics decreased during the core study and increased with both asenapine and olanzapine during the extension.

Table 5.3.

Concomitant medication

	Asenapine (n=290)	Olanzapine (n=150)
Types of medications used, n (%)		
Antidepressants		
Core study baseline	2 (0.7)	2 (1.3)
Core study endpoint	42 (14.5)	21 (14.0)
Extension endpoint	49 (16.9)	20 (13.3)
Anticholinergics		
Core study baseline	0 (0.0)	0 (0.0)
Core study endpoint	23 (7.9)	3 (2.0)
Extension endpoint	28 (9.7)	4 (2.7)
Hypnotics		
Core study baseline	21 (7.2)	12 (8.0)
Core study endpoint	21 (7.2)	6 (4.0)
Extension endpoint	31 (10.7)	7 (4.7)
Anxiolytics		
Core study baseline	61 (21.0)	31 (20.7)
Core study endpoint	46 (15.9)	19 (12.7)
Extension endpoint	61 (21.0)	25 (16.7)

Safety and Tolerability

Adverse events

Overall, fewer treatment-emergent and treatment-related AEs started during the extension than during the core study (Table 5.4). As in the core study, most AEs were mild or moderate and comparable between treatment groups. Three patients, all in the asenapine group, died during the extension. One patient, who took asenapine for 429 days, died due to acute cardiac insufficiency while on treatment. Another patient, who took asenapine for 470 days, died 4 days after the end of treatment due to disseminated arteriosclerosis and thromboembolism of pulmonary artery. The third patient, who took asenapine for 505 days, died from unknown causes 16 days after the end of treatment. The investigators considered these deaths either not related or unlikely to be related to study drug.

The most common AEs (i.e., those reported by ≥5% of patients) included insomnia, depression/depressed mood, weight increase, headache, and anxiety in both groups, as well as worsening schizophrenia for asenapine and nasopharyngitis and influenza for olanzapine. Weight increase was reported at much lower frequency during the extension than during the core study and was comparably low in both treatment groups (asenapine, 6.6%; olanzapine, 5.3%).

Body weight

At core study baseline, mean ± SD body weight was 74.7±15.4 kg in the asenapine group and 73.4±15.4 kg in the olanzapine group. Mean body weight during the extension only did not change beyond the weight gain in the core study (endpoint of core study: asenapine, 76.3±14.8 kg; olanzapine, 78.8±15.4 kg; end of entire treatment period: asenapine, 76.3±15.1 kg; olanzapine, 78.4±15.7 kg; Figure 5.3).

Extrapyramidal symptoms

Incidence rates of EPS-related AEs that started during the extension were lower for both asenapine (4.5%) and olanzapine (3.3%) compared with those that started during the core study (asenapine, 17.2%; olanzapine, 8.7%). Akathisia was the most common EPS-related AE reported in both treatment groups during the core study, but not many new cases of akathisia were reported during the extension period (Table 5.4). Changes in EPS rating scale scores from core study endpoint to extension study endpoint were equally minimal between groups, indicating that there was little change in EPS severity during the extension of treatment. No

new cases of tardive dyskinesia, a movement disorder more likely to develop with long-term antipsychotic treatment, were observed during the extension period.

Other safety and tolerability measures

No major changes in laboratory variables, including cholesterol, prolactin, and liver enzymes, and vital signs were observed in either treatment group during the extension period. The incidence of electrocardiographic-related AEs was 3.1% for asenapine and 2.7% for olanzapine during the extension only (during the entire treatment period: asenapine, 5.5%; olanzapine, 3.3%). There were no cases of QT prolongation (≥500 ms) observed in either treatment group.

Table 5.4.

Summary of adverse events

	Asenapine (n=290)		Olanzapine (n=150)	
	Started in Core Study	Started in Extension	Started in Core Study	Started in Extension
Treatment-emergent AEs, n (%)	227 (78.3)	180 (62.1)	120 (80.0)	82 (54.7)
Mild	72 (24.8)	53 (18.3)	43 (28.7)	28 (18.7)
Moderate	137 (47.2)	97 (33.4)	64 (42.7)	44 (29.3)
Severe	18 (6.2)	30 (10.3)	13 (8.7)	10 (6.7)
Treatment-related AEs, n (%)	174 (60.0)	78 (26.9)	87 (58.0)	34 (22.7)
Mild	73 (25.2)	35 (12.1)	37 (24.7)	15 (10.0)
Moderate	93 (32.1)	33 (11.4)	40 (26.7)	17 (11.3)
Severe	8 (2.8)	10 (3.4)	10 (6.7)	2 (1.3)
Serious AEs, n (%)				
Treatment emergent	26 (9.0)	54 (18.6)	6 (4.0)	12 (8.0)
Treatment related	3 (1.0)	9 (3.1)	1 (0.7)	2 (1.3)
AEs occurring in ≥5% of patients, n (%)				
Weight increase	64 (22.1)	19 (6.6)	52 (34.7)	8 (5.3)
Depression/depressed mood	55 (19.0)	23 (7.9)	21 (14.0)	10 (6.7)
Insomnia	49 (16.9)	31 (10.7)	23 (15.3)	11 (7.3)
Worsening schizophrenia symptoms	33 (11.4)	45 (15.5)	18 (12.0)	7 (4.7)
Anxiety	31 (10.7)	17 (5.9)	13 (8.7)	15 (10.0)
Headache	30 (10.3)	18 (6.2)	7 (4.7)	10 (6.7)
Akathisia	26 (9.0)	7 (2.4)	6 (4.0)	3 (2.0)
Influenza	22 (7.6)	10 (3.4)	4 (2.7)	8 (5.3)
Nasopharyngitis	22 (7.6)	5 (1.7)	6 (4.0)	9 (6.0)

AEs = adverse events.

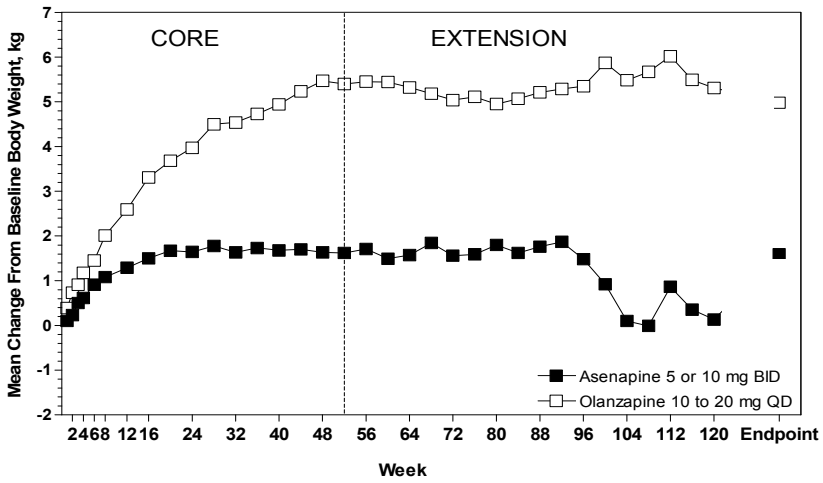


Figure 5.3.

Mean change from baseline in weight over the entire course of treatment (observed case data for all enrolled patients). BID=twice daily; QD=once daily

Efficacy

Primary and secondary efficacy measures

Figure 5.4*a* presents PANSS total score LOCF data for those subjects who continued in the extension (ITT_{ext}). Figure 5.4*b* presents LOCF data for those patients representing the total study population (ITT_{core}. In the ITT_{ext} population mean PANSS total score changes during first year of treatment were -37.0 for asenapine and -35.3 for olanzapine respectively, with a further change of 1.6 for asenapine (n=203) and -0.8 for olanzapine (n=123) at the extension study endpoint (Figure 5.4*a*). Similar to PANSS total score, no important changes in other (secondary) efficacy measures (PANSS subscales, PANSS Marder factors, CGI-S, CGI-I, and CDSS) were found. More than 80% of patients who were still taking study drug in both the asenapine and olanzapine groups had a CGI-I classification of "at least much improvement" at core study endpoint and extension endpoint.

In the OC population mean ± SD change on the PANSS total score at extension study endpoint was –35.4±17.4 with asenapine and –36.1±16.6 with olanzapine over the entire treatment period (Figure 5.4*c*). Upon truncation of the extension, clinical improvement(as reflected by a ≥20% decrease in PANSS total score from core study baseline) was reported in 92.9% of all patients still on asenapine and in 93.2% of all patients still on olanzapine;

reductions of ≥30% in PANSS total score were observed in 85.8% and 89.1% of the patients still on asenapine and olanzapine, respectively.

Health-related quality of life and functioning

During the extension and over the entire treatment period, improvement on measures of health-related quality of life (SWN, SF-12) was comparable between treatment groups among OC (Table 5.5). At endpoint of the extension, minimal additional changes in these measures were observed in either group, indicating that the improvements were maintained during the extension.

A modest association between change from baseline on clinical efficacy (PANSS total score) and functional improvement (SWN total and component scores, SF-12 physical and mental scores, and LOF total score) was found in patients treated with either asenapine or olanzapine (Figure 5.5). Equally modest intercorrelations were found between SWN total scores and SF-12 physical and mental scores (r = 0.32 and r = 0.43, respectively) and other functional outcome measures (Table 5.6).

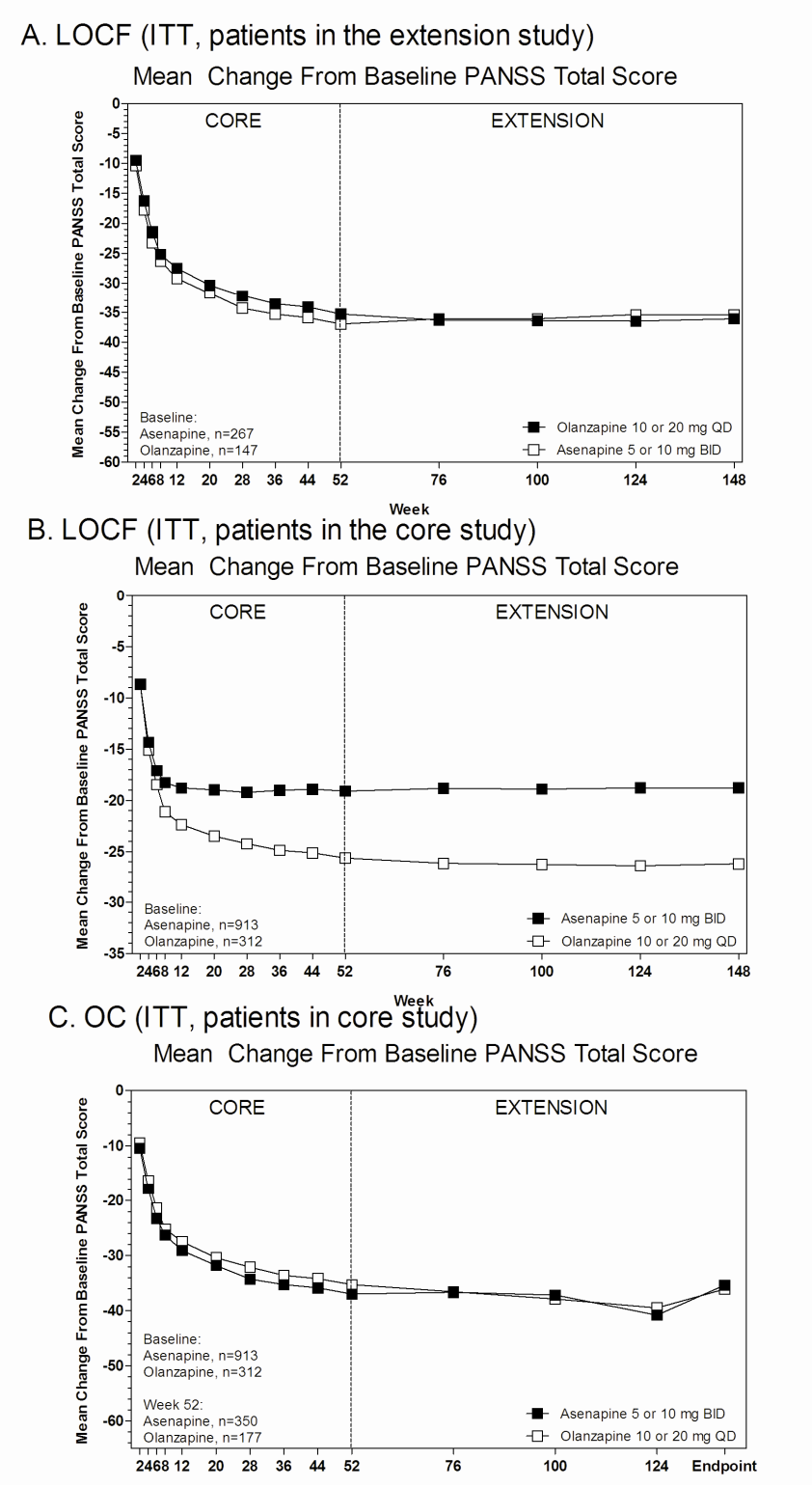


Figure 5.4

Change in Positive and Negative Syndrome Scale (PANSS) total score over the entire course of treatment.

(*a*) LOCF analysis of all patients (intent-to-treat [ITT]) entering the Extension phase.

(*b*) LOCF analysis of the Core study ITT population.

(*c*) C OC analysis of ITT patients in the core study

BID=twice daily; LOCF=last observation carried forward; OC=observed cases; QD=once daily.

Table 5.5

Summary of health-related quality-of-life measures among observed cases in the intent-to-treat population

	Asenapine (n=267)	Olanzapine (n=147)
SWN total score, mean ± SD		
Core study baseline	75.6±17.6	75.1±15.2
Change for entire treatment period	13.1±18.6	14.6±18.0
Change for extension only	–1.3±13.1	1.3±8.7
SF-12, mean ± SD		
Physical component summary		
Core study baseline	44.2±8.1	44.8±8.1
Change for entire treatment period	3.0±9.9	1.5±10.2
Change for extension only	0.2±8.0	–0.3±8.3
Mental component summary		
Core study baseline	38.1±9.6	37.0±8.5
Change for entire treatment period	6.9±12.3	8.2±12.1
Change for extension only	–0.3±9.0	0.6±7.5
LOF, n (%)		
Meets with friends ≥once per week		
Core study baseline	55 (20.6)	32 (21.9)
Core study endpoint	105 (39.3)	51 (34.7)
Extension endpoint	100 (37.5)	64 (43.5)
≥1 close relationship		
Core study baseline	42 (15.7)	22 (15.1)
Core study endpoint	87 (32.6)	41 (27.9)
Extension endpoint	87 (32.6)	46 (31.3)
Employed continuously		
Core study baseline	28 (10.5)	18 (12.3)
Core study endpoint	68 (25.5)	34 (23.1)
Extension endpoint	75 (28.1)	34 (23.1)
Very competent at work		
Core study baseline	13 (6.7)	7 (6.9)
Core study endpoint	44 (21.5)	21 (19.3)
Extension endpoint	41 (19.4)	19 (16.7)

SWN=Subjective Well-Being Under Neuroleptics; SF-12=Medical Outcomes Study 12-item Short Form; LOF=Level of Function scale.

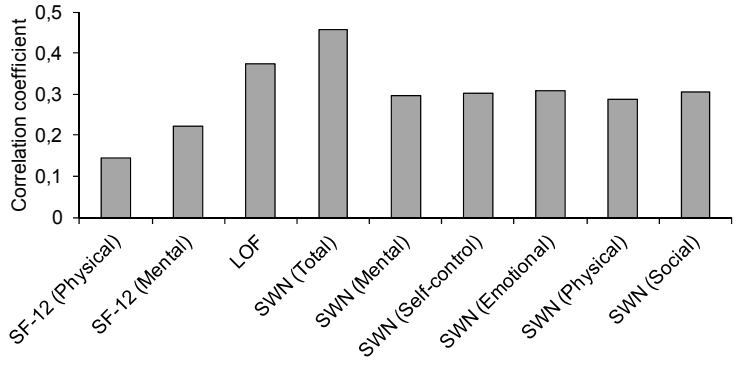


Figure 5.5

Spearmen rank correlation between change from baseline on PANSS (total score) and health-related measures (pooled treatment data). LOF=abbreviated Level of Function scale; PANSS=Positive and Negative Syndrome Scale; SF-12=Medical Outcomes Study 12-item Short Form; SWN=Subjective Well-Being Under Neuroleptics.

Table 5.6

Spearman rank correlation between change from baseline on health-related quality-of-life measures

	SF-12 Mental	SWN Total	LOF	SWN Mental	SWN Self-control	SWN Emotional	SWN Physical	SWN Social
SF-12 physical	0.04 (n=912)	0.32 (n=907)	0.12 (n=333)	0.26 (n=907)	0.23 (n=907)	0.22 (n=906)	0.31 (n=907)	0.23 (n=907)
SF-12 mental		0.43 (n=907)	0.09 (n=333)	0.33 (n=907)	0.29 (n=907)	0.38 (n=906)	0.34 (n=907)	0.37 (n=907)
SWN total			0.23 (n=336)	0.81 (n=940)	0.76 (n=940)	0.80 (n=939)	0.77 (n=940)	0.78 (n=940)
LOF				0.19 (n=336)	0.19 (n=336)	0.20 (n=335)	0.16 (n=336)	0.19 (n=336)
SWN mental					0.51 (n=940)	0.57 (n=939)	0.55 (n=940)	0.55 (n=940)
SWN self-control						0.55 (n=939)	0.52 (n-940)	0.50 (n=940)
SWN emotional							0.52 (n=939)	0.55 (n=939)
SWN physical								0.50 (n=940)

LOF=abbreviated Level of Function scale; SF-12=Medical Outcomes Study 12-item Short Form; SWN=Subjective Well-Being Under Neuroleptics.

DISCUSSION

Establishing the long-term tolerability and efficacy of drugs used to treat schizophrenia and schizoaffective disorder is important because these disorders typically require chronic, frequently life-long treatment.[14] In the initial 52-week core study in patients with schizophrenia or schizoaffective disorder, asenapine 5 or 10 mg BID and olanzapine 10 or 20 mg QD were well tolerated, with asenapine causing less weight gain than olanzapine but associated with a higher incidence of EPS.[11] Both medications improved psychotic symptoms, with PANSS total score reductions being comparable upon OC analysis but superior with olanzapine upon LOCF analysis. Consistent with the LOCF results in the core study,[11] a meta-analysis of 50 clinical trials in patients with schizophrenia or schizophrenia-like psychosis has reported that olanzapine improves PANSS total score slightly more than aripiprazole, risperidone, ziprasidone, and quetiapine, but not amisulpride or clozapine.[15] In the ITT-extension group, differences between mean changes in PANSS total score on asenapine and olanzapine appeared minimal at study endpoint over the core and extension treatment (Figures 5.4B and 5.4C). The conclusions drawn earlier for the core study are reinforced by these observations.[11]

During the extension period, no meaningful differences in weight gain and EPS were observed between treatment groups. Clinical improvement was maintained equally well with both drugs as minimal changes were observed in mean or LS mean PANSS total score. The incidence of newly emergent AEs was low on both treatments, and there were no notable changes in laboratory variables, vital signs, or electrocardiographic findings. Taken together, these observations suggest that clinical stability can be maintained with asenapine as well as olanzapine over a prolonged period of time and that AEs experienced in the early course of treatment can usually be kept under control, with few recurrences or newly emerging AEs after 1 year of treatment.

It should be acknowledged that a higher proportion of patients on asenapine discontinued the trial than on olanzapine. The difference in dropout rates between treatment groups was not only observed in the core study but also in the extension period, and may have influenced the results of the LOCF analysis that imputes for missing data, favoring the treatment group with fewer withdrawals. For a proper weighing of efficacy results over the entire treatment period (up to 3 years), the fact that double-blind treatment was truncated at an unforeseen time point

is important to consider. This so-called "humanitarian" study was open-ended (i.e., the blind for the core study was broken while patients would still be on extension treatment). Study end for the extension was delayed until patient enrollment and data cleaning for the core study was fully completed, and the database lock ready. Duration of treatment continuation was therefore of variable length. Treatment duration for each patient was also influenced by the time point of extension study enrollment, in that the first patients enrolled continued the longest time (maximally 23 months) and the last enrolled continued for the shortest time (maximally 10 days). As a result, by the end of the trial (i.e., at the moment of truncation) the majority of subjects who had once entered the extension period had only been on double-blind treatment for less than 2 years (Table 5.2).

Although the long-term data from this extension are valuable, one major limitation should be acknowledged. As explained above, no statistical analyses were done to compare results of asenapine and olanzapine. Thus, any comparative conclusions regarding the differential effects of treatment should be made with caution.

Overall, clinical improvement in the current study was associated with modest improvements in functional outcome parameters, especially with the abbreviated LOF and SWN. The relatively higher correlation between changes in PANSS total score and these scales might be because LOF and SWN (unlike SF-12) have been specifically designed to measure functional outcome in schizophrenia. It is of interest that the intercorrelations between SF-12, LOF, and SWN are relatively low, suggesting that functional outcome can be assessed in a complementary manner with these instruments.

ACKNOWLEDGMENTS

The authors thank Wim Jansen (Merck Sharp & Dohme, Oss, the Netherlands) for his meticulous review of the data presented in the manuscript. Editorial support was provided by Complete Healthcare Communications, Inc. and was supported by Merck, Whitehouse Station, NJ, USA.

We also thank the following investigators for participating in the extension: Australia: Rajan Thomas, Box Hill; Brett Emmerson, Fortitude Valley; Belgium: Claudine Mertens, Sint Denijs Westrem; Serge Seghers, Kortrijk; Eugeen De Bleeker, Sint Niklaas; Geert De

Bruecker, Lede; Antonio Gazziano, Tielt; Erik Buntinx, Alken; Joseph Peuskens, Kortenberg; Lieven De Weirdt, Sint Niklaas; Czech Republic: Marius Byss, Havlíèkùv Brod; Eva Ceskova, Brno Bohunice; Pavel Mohr, Praha; Vlasta Hanušková, Opava; Vlastimil Tichy, Praha; Eva Soukupova, Plzen; Jin Bilik, Olomouc; David Holub, Milnik; Zdenek Solle, Praha; Hana Hornochova, Brno; Iva Klanova, Sadská; France: Daniel D. Dassa, Marseille; Mocrane M. Abbar, Nîmes; Georges-Pierre Noel Badet, Dijon: Naima Kienlen, Rouffach; Daniel D. Bonnaffoux, Dole; Joel J. Thomas, Saint Nazaire; Germany: Wolfgang Maier, Bonn; Claus Normann, Freiburg; Isabella Heuser, Berlin; W. Kissling, München; Harald Freyberger, Stralsund; Poland: Mieczyslav Janiszewski, Torun; Maciej Czerwinski, Torun; Aleksander Araszkiewicz, Bydgoszcz; Ryskzard Wardenski, Zuromin; Andrzej Rajewski, Poznan; Jarslaw Strzelec, Tuszyn; Zbiniew Skurczyñski, Zabki; Irena Krupka-Matuszczyk, Katowice; Andrzej Kokoszka, Warszawa; Michal Kujawski, Zabrze; Janusz Janczewski, Chelmno; Anna Grzywa, Lublin; Ireneusz Kaczorowski, Belchatów; Przemyslaw Bogacki, Skorzewo; Russian Federation: Boris Andreyev, Leningrad; Sergey Elkin, St. Petersburg; Mikhail Popov, St. Petersburg; Mikhail Ivanov, St. Petersburg; Yury Popov, St. Petersburg, Evgeny Snedkov, St. Petersburg, Vladimir Tochilov, St. Petersburg; Vladimir Vilvanov, Chernyshevskogo; Vladislav Shamrey, St. Petersburg; Yury Aleksandrovsky, Moscow; Nikolay Neznonov, Obvodny; Margarita Morozova, Moscow; Alexander Okhapkin, Smolensk; Mikhail Sheyfer, Samara; Kausar Yakhin, Kazan; South Africa: George Hart, Johannesburg; Don Wilson, Cape Town; Robin A. Emsley, Cape Town; Hilda Russouw, Western Cape; Robert Bothwell, Cape Town; Paul Strong, Cape Town; C. Weyers, Bloemfontein; Carrllo Gagiano, Bloemfontein; P.J. Pretorius, Pretoria West; Merryl Vorster, Krugersdorp and Parktown; Spain: Jesús de la Gándara, Brugos; Manuel Angel Franco, Zamora; Miguel Angel Gonzalez Torres, Cuenca.

REFERENCES

1. Saha S, Chant D, Welham J, McGrath J. A systematic review of the prevalence of schizophrenia. *PLoS Med* 2005;2(5):e141.
2. Perala J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsa E, Pirkola S, Partonen T, Tuulio-Henriksson A, Hintikka J, Kieseppa T, Harkanen T, Koskinen S, Lonnqvist J. Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Archives of General Psychiatry* 2007;64(1):19-28.
3. Weiden PJ. EPS profiles: the atypical antipsychotics are not all the same. *Journal of Psychiatric Practice* 2007;13(1):13-24.
4. Tandon R, Jibson MD. Comparing efficacy of first-line atypical antipsychotics: no evidence of differential efficacy between risperidone, olanzapine, quetiapine, ziprasidone, and aripiprazole. *International Journal of Psychiatry in Clinical Practice* 2005;9(3):204-212.
5. Tandon R, Fleischhacker WW. Comparative efficacy of antipsychotics in the treatment of schizophrenia: a critical assessment. *Schizophrenia Research* 2005;79(2-3):145-155.
6. Volavka J, Citrome L. Oral antipsychotics for the treatment of schizophrenia: heterogeneity in efficacy and tolerability should drive decision-making. *Expert Opinion in Pharmacotherapy* 2009;10(12):1917-1928.
7. Saphris®. *asenapine sublingual tablets*. Whitehouse Station, NJ: Schering Corporation, a subsidiary of Merck & Co., Inc; 2010.
8. European Medicines Agency. Sycrest (asenapine). 2010; http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/001177/human_med_001379.jsp&murl=menus/medicines/medicines.jsp&mid=WC0b01ac058001d124&jsenabled=true. Accessed September 28, 2010.
9. Kane JM, Cohen M, Zhao J, Alphs L, Panagides J. Efficacy and safety of asenapine in a placebo- and haloperidol-controlled trial in patients with acute exacerbation of schizophrenia. *Journal of Clinical Psychopharmacology* 2010;30:1-10.
10. Potkin SG, Cohen M, Panagides J. Efficacy and tolerability of asenapine in acute schizophrenia: a placebo- and risperidone-controlled trial. *Journal of Clinical Psychiatry* 2007;68(10):1492-1500.
11. Schoemaker J, Naber D, Vrijland P, Panagides J, Emsley R. Long-term assessment of Asenapine vs. Olanzapine in patients with schizophrenia or schizoaffective disorder. *Pharmacopsychiatry* 2010;43(4):138-146.
12. Strauss JS, Carpenter WT, Jr. The prediction of outcome in schizophrenia. I. Characteristics of outcome. *Archives of General Psychiatry* 1972;27(6):739-746.
13. Strauss JS, Carpenter WT, Jr. Prediction of outcome in schizophrenia. III. Five-year outcome and its predictors. *Archives of General Psychiatry* 1977;34(2):159-163.
14. Agid O, Kapur S, Remington G. Emerging drugs for schizophrenia. *Expert Opinion on Emerging Drugs* 2008;13(3):479-495.
15. Komossa K, Rummel-Kluge C, Hunger H, Schmid F, Schwarz S, Duggan L, Kissling W, Leucht S. Olanzapine versus other atypical antipsychotics for schizophrenia. *Cochrane Database of Systematic Reviews* 2010;3:CD006654.

Chapter 6

Evaluation of satisfaction with treatment and drop-out in schizophrenia

Towards a better understanding of factors reducing power in randomized, controlled trials (II) [1]

[1] This chapter was published as:

Joep H. Schoemaker, Ad J.J.M. Vingerhoets, Robin A. Emsley. (2018).

Factors associated with satisfaction with treatment and trial discontinuation in chronic schizophrenia.

CNS Spectrums.

DOI: http:dx.doi.org/10. 1017/S109285291700044X.

ABSTRACT

Introduction: Despite consistently high discontinuation rates due to the withdrawal of consent (WOC) and insufficient therapeutic effect (ITE) in schizophrenia trials, insight into the underlying factors contributing to poor satisfaction with treatment and drop-out is limited. A better understanding of these factors could help to improve trial design and completion rates.

Methods: Using data from 1136 trial participants with schizophrenia or schizoaffective disorder, we explored associations between predictor variables with (1) drop-out due to WOC and ITE, and (2) satisfaction with treatment among patients and investigators by means of hierarchic multiple regression analyses.

Results: ITE was associated with poor clinical improvement, investigator's satisfaction with treatment, and patient's insight into their own disease, whereas WOC only showed a meaningful association with a poor patient's satisfaction with treatment. Investigator satisfaction with treatment appeared most strongly associated with PANSS positive factor endpoint scores, whereas patient satisfaction with treatment was best predicted by the endpoint score on the PANSS emotional distress factor. The occurrence of severe side effects showed no meaningful association with satisfaction with treatment among investigators and patients, neither did a patient's experienced psychopathology, or self-rating of functional impairment.

Conclusion: Whereas trial discontinuation due to ITE is associated with poor treatment effectiveness, a patient's decision to withdraw from an antipsychotic trial remains unpredictable and may occur even when the investigator observes a global clinical improvement and is satisfied with treatment.

INTRODUCTION

Overall attrition rates in randomized, controlled trials (RCT) in schizophrenia frequently exceed 50% and appeared to increase with 1% with each publication year in the period from 1960 to 2000.[1] As a result, trials have become increasingly inefficient, and sample sizes need to correct for this in order to preserve sufficient power. Better knowledge and understanding of the underlying factors contributing to drop-out (e.g., baseline characteristics, trial procedures, and treatment effects) could be helpful to improve clinical trial design and completion rates.

Patients' withdrawal of consent (WOC) and investigator-rated insufficient therapeutic effect (ITE) are commonly observed in RCT with antipsychotics. In several large-scale, long-term, antipsychotic RCT conducted during the last decade, drop-out rates due to WOC varied between 29% and 40%.[2-5] Despite a consistently high percentage of drop-out due to WOC among patients with schizophrenia participating in RCT, the underlying reasons causing a participant to withdraw from a clinical study have, to the best of our knowledge, never been systematically investigated. Can WOC perhaps be associated with, and regarded as a form of poor compliance that is so commonly seen in schizophrenia patients? In daily practice, treatment of schizophrenia is often challenged by various degrees of treatment disengagement or non-adherence.[6] Adherence rates in schizophrenia are substantially lower than the adherence rates in other, chronic (and potentially life-threatening) conditions. The median non-adherence rates to antipsychotics typically range from 40 to 55%[7,8] and lack of insight and poor satisfaction with treatment have repeatedly been identified as major determinants.[8-12] However, also subjective well-being has been shown to exert a separate, independent influence on compliance, irrespective of the presence of clinical symptoms.[13-15] In contrast, the severity of baseline psychotic symptoms, illness duration, the presence of mood symptoms (or diagnosis of schizoaffective disorder), inpatient status, history of substance abuse, and socio-demographic variables have yielded mixed results regarding associations with non-adherence.[12]

Poor treatment response, leading to drop-out, occurs on average around 20% in long-term RCTs in schizophrenia.[2,3,16] Upon study discontinuation, investigators are usually requested to record the single most important reason for drop-out. This can be either lack of efficacy, adverse events (AEs), WOC, or other reasons such as failure to return for follow-up visits.

This raises the question, what makes a patient decide to withdraw from a trial when investigators are satisfied with the efficacy and tolerability of the treatment. It is also of interest to know the extent to which trial discontinuation due to WOC and ITE is specifically associated with a poor satisfaction with treatment of patients (PST) and of investigators (IST).

Factors earlier described as being possibly associated with patients' satisfaction with treatment included measures of 'distress,' 'subjective well-being,' and 'functional outcome'.[17-20] Several longitudinal studies reported a positive association between the amelioration of depressive symptoms and a patient's subjective well-being or quality of life.[21-24] Interestingly, in standardized interviews, patients diagnosed with schizophrenia did not highly rank depressive thoughts and emotions as a priority treatment goal, whereas physicians particularly attached value to improved cognitive abilities and reduction of disease-related symptoms.[18] Although factors determining satisfaction with treatment have become a research topic of growing interest, the majority of studies in this area so far have considerable limitations such as a small sample size, open-label treatment, and weighted selection of respondents towards those who have good experiences in survey analysis.[17,25-27]

In an attempt to identify determinants of drop-out in RCTs, we extracted baseline data and treatment results from a large-scale, multiregional RCT for a posthoc analysis. We hypothesized, that drop-out due to ITE is primarily the result and reflection of a clinician's satisfaction with treatment and decision to discontinue, whereas drop-out due to WOC is primarily the result and reflection of a patient's satisfaction with treatment and decision to discontinue. We further hypothesized that IST and PST are driven by partly different treatment effects and not necessarily associated in the case of drop-out.

METHODS

A 52-week, double-blind, randomized, active-controlled, two-armed, multi-regional study was designed to explore the long-term efficacy and safety of asenapine in comparison with olanzapine in a large sample of patients diagnosed with schizophrenia or schizoaffective disorder (registered under number NCT00212784 at ClinicalTrials.gov). The posthoc analysis reported here relies on the baseline data and treatment effects from all participants in this

study, except those who had never received treatment with antipsychotics before. Details of the underlying study design and entry criteria have been described elsewhere.[28]

Study participants

Twelve-hundred-twenty-five in-patients and out-patients fulfilling criteria for schizophrenia (SCZ) or schizoaffective disorder (SAD) according to the Diagnostic and Statistical Manual of Mental Disorders (4th ed, text rev.; DSM-IV-TR, American Psychiatric Association, 2000) were enrolled across 102 sites in Europe, Russia, Australia, and South Africa. They all received up to 52 weeks double-blind, double-dummy treatment with asenapine (5-10 mg bid) or olanzapine (10-20 mg qd).[29] Eighty-three of them had never been on antipsychotics before and were excluded from the current analysis because satisfaction with treatment was measured in comparison with treatment received before (see below). From the remaining 1142 enrolled patients with a non-first episode of schizophrenia or schizoaffective disorder, six were not further treated after screening, leaving a total of 1136 participants for evaluation.

Outcome measures

For all study participants not completing the 52 weeks double-blind treatment, investigators were instructed to record the single most important reason for drop-out, selecting from the following list: (*1*) adverse event or serious adverse event (including hospitalization related to worsening of disease), (*2*) lack of efficacy, (*3*) lost to follow-up, (*4*) withdrawal of consent, or (*5*) other reason. Patients who dropped out due to lack of efficacy or due to the serious adverse event 'hospitalization, related to worsening of disease' were considered as one group (i.e., dropping out due to ITE). Asymmetry in the data did not allow considering patients lost to follow-up (LFU) and patients withdrawing consent (WOC as one group, with PANSS total scores showing a mean reduction (at the last assessment prior to drop-out) of approximately 30 points in the LFU group versus 14 in the WOC group and less than 1 points in patients dropping out due to lack of efficacy or hospitalization due to worsening. The WOC group, therefore, consisted only of patients who withdrew consent.

For the assessment of satisfaction with treatment (PST and IST), patients and investigators were asked to rate their satisfaction with treatment in comparison with previous medication administered for their disease on a five-point Likert scale, ranging from (1) *much worse* to (5) *much better.*

Drop-out due to WOC and drop-out due to ITE were used as dependent variables in multiple regression models, exploring the extent to which these specific outcomes could be predicted by particular baseline characteristics and treatment results. In a subsequent step, using similar multiple regression models, the influence of specific symptom dimensions, subjective well-being, and health impairment on satisfaction with treatment among patients and investigators was explored.

Outcome predictors

To explore the influence of patient characteristics on outcome, the following baseline data were collected: gender, level of education (# years), marital status (with or without a partner), recent substance abuse (graded as abstinence, use without impairment, abuse, or dependence), diagnostic subtype (SCZ or SAD), number of years symptoms were present, hospitalization status (in- or out-patient), disease severity, and randomized treatment received. We additionally collected individual patient data for factors earlier described to be associated with drop-out due to lack of efficacy or treatment disengagement, including clinical improvement, level of insight, occurrence of side effects (as reflected by the total number of severe adverse events with possible, probable, or definitive relationship with treatment according to the investigator), satisfaction with treatment, and subjective well-being.[8-15]

Variables were assessed and considered to collectively predict outcome in the current analysis as follows. Disease severity and change from baseline were rated by the clinician on the Clinical Global Impression – Severity (CGI-S) and - Improvement (CGI-I) subscales.[30] Together with IST scores and the occurrence of disturbing side effects, CGI-I ratings were considered as indices of clinician-rated treatment response, or '*effectiveness*.' Individual symptom severity was rated by clinicians on the 30-item PANSS.[31] The item scores were clustered into five dimensions of core disease symptoms following Kelly et al. [32]: (*1*) the negative component (including the items 'conceptual disorganization', 'blunted affect', 'emotional withdrawal', 'poor rapport', 'social apathy', 'lack of spontaneity', 'motor retardation', 'active social avoidance'); (*2*) the positive component ('delusions', 'hallucinations', 'grandiosity', 'suspiciousness', 'unusual thought content'); (*3*) the disorganized component ('conceptual disorganization', 'stereotyped thinking', 'mannerism and posturing', 'disorientation', 'poor attention', 'lack of insight', 'preoccupation'); (*4*) the excited component ('excitement', 'hostility', 'uncooperativeness', 'poor impulse control');

and (*5*) the emotional distress component ('anxiety', 'guilt', 'tension', 'depression'). Cluster scores on these components at endpoint, together with the occurrence of disturbing side effects, were considered as indices of clinician-observed psychopathology, or '*illness*.' The clinician-rated PANSS cluster scores and patient self-ratings on two separate instruments were used to confirm the earlier observations that patients attach more value to the amelioration of impairment and the allevaiation of distress, whereas investigators attach more value to the amelioration of cognitive abilities and positive symptoms.[17-20] Scores on PANSS item G12 'lack of judgment and insight' and PST were used as indices of a patient's '*treatment engagement*.' Patient self-ratings on the Subjective Well-being under Neuroleptic treatment (SWN) scale and the physical and mental component scales of the Medical Outcomes Study 12-Item Short Form (SF-12) health survey were considered as indices of disease as experienced by the patient, or '*impairment*'.[33,34] Above described indices of 'effectiveness' and 'treatment engagement' were entered into the regression models predicting drop-out, whereas the indices of 'illness' and 'impairment' were entered into the regression models predicting satisfaction with treatment.

Compliance was assessed by the clinician through pill-counts of returned medication at each follow-up visit and ranked as 'excellent' when deviating not more than 5%, 'satisfactory' when deviating not more than 25%, and 'poor' when deviating more than 25% of prescribed dosages over the entire treatment period. Although it seemed to make sense to include measurements of compliance as an index of 'treatment engagement' in the regression models, since clinicians judged compliance satisfactory in more than 95% of cases and excellent in more than 70% of cases, these rankings were not considered informative enough to include them as possible predictors of any of the outcome variables.

Preselected predictor variables were tested for independence and relevance through correlational analysis, whereby a Pearson correlation coefficient between variables of 0.25 or higher (at a statistical significance level of $p<0.05$) was considered to be meaningful. Predictor variables showing lowest correlation with other predictor variables and highest correlation with one or more of the outcome variables (WOC, ITE, PST, or IST) were preserved for the regression models. Thus, a trade-off was made between available indicators of medical history (e.g., age vs. years symptoms present, alcohol vs. substance abuse), efficacy (e.g., CGI-I vs. PANSS total score change from baseline, treatment duration, percentage of treatment period hospitalized, or hospitalized at last day of treatment), and

safety (e.g., related, severe vs. serious or moderate AEs, AEs of any intensity or relationship with treatment, weight gain, or extrapyramidal symptoms). All analyses were carried out using SPSS (Windows) Version 23.0.

Regression models and analyses

The independent (baseline and treatment response) variables were split into blocks and entered in a hierarchic multiple regression procedure according to the chronological order in which they became available. Adjusted R^2 was used as a measure of determination of the models, and change in R^2 as a measure of improvement obtained by adding the variables of the subsequent blocks. Sociodemographics (gender, education, marital status) were entered first, followed by details of medical history (substance abuse, DSM-IV diagnosis, disease duration, hospitalization status), baseline severity (CGI-S), and treatment received in steps two to four. As there were no meaningful predictors identified in the analysis of the first four models (containing demographics, medical history, baseline severity, and treatment), blocks one to four were combined into one in the final regression analyses and referred to as model I comprising all 'risk factors.' In subsequent steps, variables measuring 'effectiveness' and 'engagement,' or 'illness' and 'impairment' were entered as predictors of drop-out (WOC, ITE) or satisfaction with treatment (PST, IST) respectively, and referred to as models II and III. Missing values were excluded pair-wise, and preliminary analyses were conducted to ensure normality and linearity of the data, as well as the absence of interfering multicollinearity.

RESULTS

Baseline characteristics, treatment response, and outcome

Descriptive statistics and intercorrelations of baseline characteristics are presented in Table 1. Mean total PANSS score at baseline was 92 (range 50-146), indicative of a population of markedly ill patients.[35] A minority of the participants was diagnosed with schizoaffective disorder (SAD=159, 14.0%) and the great majority with schizophrenia (SCZ=977, 86.0%). A total of 491 participants (43.2%) completed one-year double-blind treatment. As shown in Fig. 6.1, overall attrition was 20.0% during the first six weeks, gradually increasing over time

to 56.8% after 52 weeks of treatment. Main reasons for study discontinuation over the entire treatment period were ITE (n=260, 22.9%) and WOC (n=235, 20.7%).

Descriptive statistics and intercorrelations of efficacy and safety parameters are presented in Table 6.2. CGI-I endpoint scores indicate that 45% percent of the participants were considered to have at least minimally improved at endpoint, 23% to have minimally worsened or not changed, and 10% to have worsened much or very much since the onset of their treatment. The majority of participants (67.1%) experienced no, or only mild adverse events, whereas 78 participants (6.9%) experienced one or more severe adverse event, at least possibly related to treatment (AE). In 156 cases, PST or IST scores were not available, leaving 980 participants for a cross-comparison of satisfaction with treatment.

Whereas there were no meaningful intercorrelations between baseline characteristics, a moderate correlation between the SF-12 mental component and PANSS emotional factor scores (r = -0.37, $p<0.01$) was observed. Also, SWN total scores showed a moderate correlation with CGI and PANSS factor scores, and fairly high correlation with SF-12 physical (r = 0.40) and SF-12 mental (r = 0.46) component scores. Intercorrelations between CGI and PANSS factor scores were substantial to very high (range r = 0.47-0.80).

Factors associated with trial discontinuation

The results of the hierarchical multiple regression analyses are summarized in Tables 6.3*a* and 6.3*b*. The extent to which drop-out due to ITE could be predicted on the basis of treatment response was moderate (R^2 = 0.383). A relatively high baseline severity and lack of improvement on CGI, as well as the lack of insight at endpoint (PANSS item G12) and a poor investigator's satisfaction with treatment, were found to be significant predictors of study discontinuation due to ITE. Drop-out due to ITE appeared to be positively associated with a patient's satisfaction with treatment (standardized β = 0.144, $p<0.001$).

The extent to which drop-out due to WOC could be predicted was very modest (R^2 = 0.112). Mild illness at baseline and a poor patient satisfaction with treatment were identified as significant predictors (albeit with a very weak power). Drop-out due to WOC appeared to be positively associated with an investigator's satisfaction with treatment (standardized β = 0.260, $p<0.001$). The occurrence and frequency of severe side effects were not predictive for drop-out due to ITE or WOC.

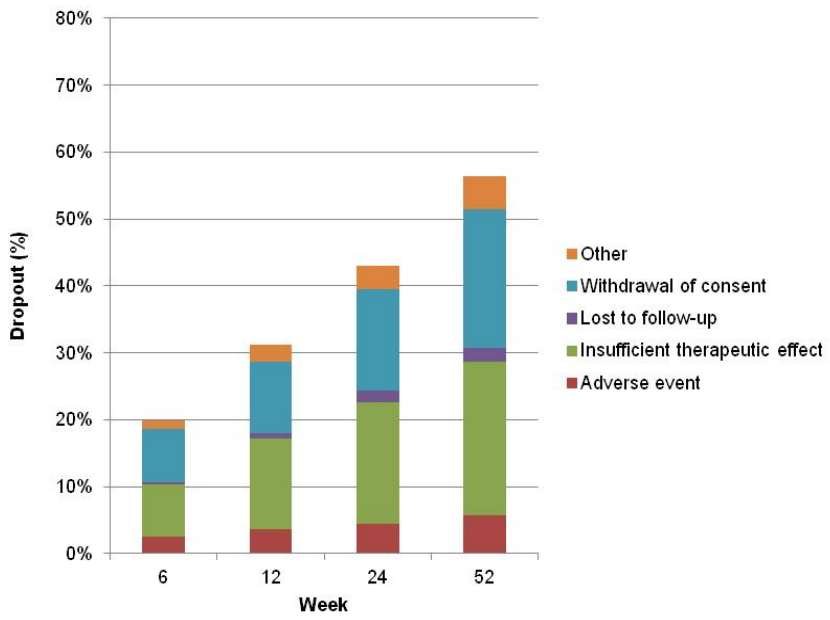


Fig 6.1

Drop-out (%) over time (weeks) by reason (insufficient therapeutic effect, withdrawal of consent, adverse event, lost to follow-up, or other)

Factors associated with satisfaction with treatment

Almost 60% of the variance in IST and more than 40% of the variance in PST could be explained by by the full model (steps I-III). Occurrence of severe side effects was only marginally associated with IST (standardized β = -0.078, p=0.001) and not at all with PST (standardized β = -0.040, p=0.127). High scores on the various PANSS factors predicted poor satisfaction with treatment, although the disorganized factor was found to have a modest, but significant effect on IST only (standardized β = -0.107, p=0.032) and not on PST (standardized β = -0.070, p=0.224). The PANSS positive factor had the strongest association with IST (standardized β = -0.293, p<0.001), whereas PST was more strongly predicted by the PANSS emotional distress factor (standardized β = -0.267, p<0.001). A relatively mild illness at baseline and diagnosis of SAD contributed negatively to satisfaction with treatment, although these two variables together only explained around 7% of the variance in IST and around 3% of the variance in PST. Patient self-ratings did not show any significant effect on IST, but the SF-12 Physical and Mental component scores appeared a significant predictor for PST (although only explaining 1% of the variance when the other variables had been accounted for).

DISCUSSION

The aim of the present study was to identify major determinants of trial discontinuation in a multiregional RCT. We hypothesized that patients and investigators attach different value to the effects of treatment and that poor patient satisfaction is the leading cause of WOC, whereas poor investigator satisfaction is the leading cause of ITE. Results were in line with our hypotheses: drop-out due to ITE and WOC showed strongest association with IST and PST respectively, whereby the PANSS positive and excited factor scores at endpoint were the main determinants of IST, and the PANSS emotional factor score the main determinant of PST.

As expected, drop-out due to ITE was most strongly associated with poor treatment effectiveness. In contrast, drop-out due to WOC appeared difficult to predict, and was not meaningfully associated with attenuated clinical improvement, the occurrence of severe side effects, or a patient's lack of insight. Considering WOC as an ultimate form of treatment non-adherence, our results failed to corroborate the findings of naturalistic studies that lack of insight has a negative impact on drug compliance or treatment adherence.[36,37] Neither do our results confirm that undesirable side effects are a major contributor to WOC or PST, as could have been expected on the basis of surveys held among patients and psychiatrists.[26,38,39] The present findings are also not in line with the earlier observations of Perkins et al. (2008), showing that poor treatment efficacy and tolerability are predictors of poor medication adherence, and that both - in combination – are associated with an increased likelihood of discontinuation against medical advice.[40] Both Perkin's and our study entailed secondary analyses of a flexible-dose, 52-week RCT with antipsychotics in patients diagnosed (according to DSM-IV criteria) with schizophrenia or schizoaffective disorder, concomitantly allowing for adjunctive medication such as benzodiazepines and anticholinergics. Main differences between the two RCTs were that in Perkins's study the predictor analysis was done in first-episode patients (excluded from our analysis), a lower percentage of participants completed one-year treatment (29.8% vs. 43.2%), and more participants withdrew consent (28.8% vs. 20.7%). Adherence to prescribed drug intake was 50% to 75% in Perkins study, and presumably lower than the adherence rates observed in our study.

Although differences in completion rates and adherence between the two studies may have been due to chance, it could also well be that these reflect an underlying difference in

treatment effects or expectations among first-episode patients compared to more chronic patients. It is intriguing that, although drop-out due to ITE was positively associated with poor treatment effectiveness, there was a remarkable trend visible in our data for a negative association between drop-out due to WOC and CGI-I, suggesting a tendency towards *improvement* rather than worsening among participants withdrawing consent. This is in line with the recognition by psychiatrists that patients feeling better and thinking that their medication is no longer necessary are important causes of medication discontinuation in schizophrenia.[38]

IST and PST were not influenced to a similar extent by the perceived or experienced severity of symptoms in specific domains. Whereas IST was most strongly associated with severity of 'positive symptoms' at discontinuation and (to a lesser extent) side effects, PST was predicted best by the severity of 'emotional distress' and, to a lesser extent, experienced impairment in functioning. The presence of mood symptoms, as reflected by a diagnosis of SAD, had a modest negative effect on both IST and PST.

These results are in line with earlier findings demonstrating that investigators and patients weigh the merits of antipsychotic treatment in partially different ways. For example, Fervaha and co-workers found that change in overall illness severity, as determined by clinicians, was not interchangeable with patients' view of improvement of their illness status. In their study, change in positive psychotic symptoms was the strongest predictor of clinician-rated illness severity scores, whereas improvement in depressive symptoms was the strongest predictor of improvement in illness severity, as rated by the patient.[41,42] These findings, together with our results, concord with the interview findings of Kuhnigk et al., demonstrating that clinicians are primarily focused on psychotic symptom control while patients diagnosed with schizophrenia rank fewer depressive thoughts and emotional distress of highest importance as a treatment goal.[18]

The strengths of our study are the relatively large sample size, the inclusion of diverse populations from multiple regions in which the RCT was executed, the reliable investigator ratings, and the relatively high level of compliance. More than one thousand patients were enrolled across three continents. There may have been various reasons for patients to participate, including economic (health insurance) reasons, poor response to previous treatment, and social reasons, each having a potential impact on overall motivation to stay in the trial. Region-specific differences in the main reason to participate may have been

mitigated through the wide sample of patients enrolled. Although the assessment of adherence through pill-counts is not always reliable and may underestimate medication non-adherence, compliance was satisfactory in the large majority of participants and substantially better than could normally be expected on the basis of commonly reported poor adherence rates to prescribed antipsychotics.[7,8,43] Last but not least, all investigators were required to participate in an inter-rater training program before trial execution, ensuring all PANSS items were evaluated according to pre-defined criteria.

There are also several limitations inherent to our approach: (*1*) The original study was not designed to demonstrate the disparity in satisfaction with treatment among investigators and patients. Neither were investigators required to justify recorded reasons for discontinuation or trained to base their choice of primary reason on objective criteria. (*2*) Although results seem to confirm the validity of our assumption that poor patient satisfaction is the leading cause of WOC and poor investigator satisfaction is the leading cause of ITE, they have to be interpreted with caution since our evaluation method did not allow us to discriminate between cause and effect. More precisely, satisfaction with treatment was assessed *after* discontinuation, but a patient's and/or investigator's disappointment may well have *preceded* the decision to discontinue and even may have been its prime reason (regardless of overall clinical effectiveness). (*3*) IST and PST were used as independent variables in the predictive models for drop-out, although both showed substantial overlap in the overall dataset. The same variables were also used as dependent variables in the predictive models for satisfaction with treatment, although we did not correct for multiplicity. (*4*) Regarding the use of PANSS factors as indicators of symptom severity in the different domains, it must be noted that underlying symptom clusters were derived from a single scale, designed to assess the severity of illness in its entirety. This may explain the relatively strong associations between factor scores so that results obtained should be considered as merely indicative and requiring confirmation in follow-up research. (*5*) Since participants in RCT may strongly differ from an epidemiologic sample, results of this RCT may not be generalized to the average practice population of patients suffering from schizophrenia. (*6*) Extrapyramidal side effects and weight gain have repeatedly been identified as contributors to antipsychotic discontinuation.[44] Although separate scales for the assessment of depressive symptoms, social functioning, extrapyramidal symptoms, and weight gain had been used in the present study, the overall flat ratings on these measures did not allow us to include these as possible predictors in the

regression models. Instead, we included the more generic PANSS factor scores and occurrence of severe AEs as main measures of psychopathology and tolerability. Our results can therefore not rule out, that specific symptoms (such as akathisia or weight gain) strongly influenced satisfaction with treatment and drop-out. (*7*) Use of SWN and SF-12 as measures of impairment has its limitations because it may prove difficult to obtain completed forms or reliable entries on these in the case of severe psychopathology. The relatively high variance in endpoint scores on these self-rated instruments (even among the almost five hundred participants completing the study) may be a reflection of that difficulty and could well have been prohibitive for identifying SWN or SF-12 scores as strong predictors of satisfaction with treatment.

In conclusion, our findings suggest that poor investigator satisfaction with treatment is the leading cause of an investigator's decision to discontinue medication, and additionally is closely related to poor treatment effectiveness, whereas a patient's decision to withdraw from an antipsychotic trial, although slightly associated with poor patient satisfaction with treatment, is rather unpredictable on the basis of the clinician-rated clinical improvement. Emotional distress appears to have a relatively strong impact on patients' satisfaction with treatment, so that close monitoring of, and adequate measures to mitigate stress factors might somewhat reduce the chance for patients to withdraw consent. These measures could include efforts by the clinician to (*1*) foster a positive relationship with patients, encouraging patients to continue treatment as soon as improvement occurs, (*2*) promptly identify depressive and anxious feelings at onset, and (*3*) provide supportive treatment for these symptoms if necessary. Family members and caregivers may also be actively involved in the plan of care, ensuring a tension-free environment in which the patient lives, and adequately supporting a patient whenever habituation to new circumstances is necessary.

Ethical standards

All participants provided written informed consent for the collection, processing, analysis, and reporting of their anonymous medical information as necessary for scientific purposes, including use in future medical or pharmaceutical research, prior to their inclusion in the study. The RCT was approved by the appropriate ethics committees and therefore performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

REFERENCES

1. Wahlbeck K, Tuunainen A, Ahokas A, Leucht S. Dropout rates in randomised antipsychotic drug trials. *Psychopharmacology* 2001;155(3):230-233.
2. Lieberman JA, Stroup TS, McEvoy JP, Swartz MS, Rosenheck RA, Perkins DO, Keefe RS, Davis SM, Davis CE, Lebowitz BD. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England Journal of Medicine* 2005;353(12):1209-1223.
3. Kahn RS, Fleischhacker WW, Boter H, Davidson M, Vergouwe Y, Keet IP, Gheorghe MD, Rybakowski JK, Galderisi S, Libiger J. Effectiveness of antipsychotic drugs in first-episode schizophrenia and schizophreniform disorder: an open randomised clinical trial. *Lancet* 2008;371(9618):1085-1097.
4. Liu-Seifert H, Adams DH, Kinon BJ. Discontinuation of treatment of schizophrenic patients is driven by poor symptom response: a pooled post-hoc analysis of four atypical antipsychotic drugs. *BMC Medicine* 2005;3(1):1-10.
5. McEvoy JP, Lieberman JA, Perkins DO, Hamer RM, Gu H, Lazarus A, Sweitzer D, Olexy C, Weiden P, Strakowski SD. Efficacy and tolerability of olanzapine, quetiapine, and risperidone in the treatment of early psychosis: a randomized, double-blind 52-week comparison. *American Journal of Psychiatry* 2007;164:1050-1060.
6. Mert DG, Turgut NH, Kelleci M, Semiz M. Perspectives on reasons of medication nonadherence in psychiatric patients. *Patient Preference and Adherence* 2015;9:87-93.
7. Young JL, Zonana HV, Shepler L. Medication noncompliance in schizophrenia: codification and update. *Journal of the American Academy of Psychiatry* 1986;14(2):105-122.
8. Fenton WS, Blyler CR, Heinssen RK. Determinants of medication compliance in schizophrenia: empirical and clinical findings. *Schizophrenia Bulletin* 1997;23(4):637-651.
9. Cramer JA, Rosenheck R. Compliance with medication regimens for mental and physical disorders. *Psychiatric Services* 1998;49(2):196-201.
10. Lacro JP, Dunn LB, Dolder CR, Jeste DV. Prevalence of and Risk Factors for Medication Nonadherence in Patients With Schizophrenia: A Comprehensive Review of Recent Literature.[CME]. *Journal of Clinical Psychiatry* 2002;63(10):892-909.
11. Higashi K, Medic G, Littlewood KJ, Diez T, Granström O, De Hert M. Medication adherence in schizophrenia: factors influencing adherence and consequences of nonadherence, a systematic literature review. *Therapeutic Advances in Psychopharmacology* 2013;3(4):200-218.
12. Cabeza IGa, Amador MS, Lopez CA, de Chavez MG. Subjective response to antipsychotics in schizophrenic patients: clinical implications and related factors. *Schizophrenia Research* 2000;41(2):349-355.
13. Marder S. Subjective experiences on antipsychotic medications: synthesis and conclusions. *Acta Psychiatrica Scandinavica* 2005;111(s427):43-46.
14. Karow A, Czekalla J, Dittmann RW, Schacht A, Wagner T, Lambert M, Schimmelmann BG, Naber D. Association of subjective well-being, symptoms, and side effects with compliance after 12 months of treatment in schizophrenia. *Journal of Clinical Psychiatry* 2007;68(1):75-80.
15. Lambert M, Schimmelmann BG, Naber D, Schacht A, Karow A, Wagner T, Czekalla J. Prediction of remission as a combination of symptomatic and functional remission and adequate subjective well-being in 2960 patients with schizophrenia. *Journal of Clinical Psychiatry* 2006;67(11):1690-1697.
16. Haro JM, Suarez D, Novick D, Brown J, Usall J, Naber D, Group SS. Three-year antipsychotic effectiveness in the outpatient care of schizophrenia: observational versus randomized studies results. *European Neuropsychopharmacology* 2007;17(4):235-244.
17. Chue P. The relationship between patient satisfaction and treatment outcomes in schizophrenia. *Journal of Psychopharmacology* 2006;20(6 Suppl):38-56.
18. Kuhnigk O, Slawik L, Meyer J, Naber D, Reimer J. Valuation and attainment of treatment goals in schizophrenia: perspectives of patients, relatives, physicians, and payers. *Journal of Psychiatric Practice* 2012;18(5):321-328.
19. Shumway M, Saunders T, Shern D, Pines E, Downs A, Burbine T, Beller J. Preferences for schizophrenia treatment outcomes among public policy makers, consumers, families, and providers. *Psychiatric Services* 2003:1124-1128.
20. Hofer A, Rettenbacher MA, Widschwendter CG, Kemmler G, Hummer M, Fleischhacker WW. Correlates of subjective and functional outcomes in outpatient clinic attendees with schizophrenia and schizoaffective disorder. *European Archives of Psychiatry and Clinical Neurosciences* 2006;256(4):246-255.

21. Kim JH, Ann JH, Kim MJ. Relationship between improvements of subjective well-being and depressive symptoms during acute treatment of schizophrenia with atypical antipsychotics. *Journal of Clinical Pharmacy and Therapeutics* 2011;36(2):172-178.
22. Narvaez JM, Twamley EW, McKibbin CL, Heaton RK, Patterson TL. Subjective and objective quality of life in schizophrenia. *Schizophrenia Research* 2008;98(1-3):201-208.
23. Karow A, Moritz S, Lambert M, Schoder S, Krausz M. PANSS syndromes and quality of life in schizophrenia. *Psychopathology* 2005;38(6):320-326.
24. Fitzgerald PB, de Castella ARA, Filia K, Collins J, Brewer K, Williams CL, Davey P, Kulkarni J. A longitudinal study of patient- and observer-rated quality of life in schizophrenia. *Psychiatry Research* 2003;119(1-2):55-62.
25. Hellewell JS. Patients' subjective experiences of antipsychotics. *CNS Drugs* 2002;16(7):457-471.
26. Fakhoury WK, Wright D, Wallace M. Prevalence and extent of distress of adverse effects of antipsychotics among callers to a United Kingdom National Mental Health Helpline. *International Clinical Psychopharmacology* 2001;16(3):153-162.
27. Kohler S, Unger T, Hoffmann S, Steinacher B, Fydrich T. Patient satisfaction with inpatient psychiatric treatment and its relation to treatment outcome in unipolar depression and schizophrenia. *International Journal of Psychiatry in Clinical Practice* 2015;19(2):119-123.
28. Schoemaker J, Naber D, Vrijland P, Panagides J, Emsley R. Long-term assessment of Asenapine vs. Olanzapine in patients with schizophrenia or schizoaffective disorder. *Pharmacopsychiatry* 2010;43(4):138-146.
29. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders DSM-IV-TR fourth edition (text revision).* Washington, DC: American Psychiatric Association; 2000.
30. Guy W. Clinical global impression scale. *The ECDEU Assessment Manual for Psychopharmacology-Revised Volume DHEW Publ No ADM* 1976;76(338):218-222.
31. Kay SR, Flszbein A, Opfer LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 1987;13(2):261-276.
32. Kelley ME, White L, Compton MT, Harvey PD. Subscale structure for the Positive and Negative Syndrome Scale (PANSS): A proposed solution focused on clinical validity. *Psychiatry Research* 2013;205(1):137-142.
33. Naber D, Moritz S, Lambert M, Rajonk F, Holzbach R, Mass R, Andresen B, Frank P, Rüdiger H, Reinhard M. Improvement of schizophrenic patients' subjective well-being under atypical antipsychotic drugs. *Schizophrenia Research* 2001;50(1):79-88.
34. Ware Jr JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care* 1996;34(3):220-233.
35. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophrenia Research* 2005;79(2):231-238.
36. Dassa D, Boyer L, Benoit M, Bourcet S, Raymondet P, Bottai T. Factors associated with medication non-adherence in patients suffering from schizophrenia: a cross-sectional study in a universal coverage health-care system. *Australian & New Zealand Journal of Psychiatry* 2010;44(10):921-928.
37. Ascher-Svanum H, Faries DE, Zhu B, Ernst FR, Swanson JW. Medication adherence and long-term functional outcomes in the treatment of schizophrenia in usual care. *Journal of Clinical Psychiatry* 2006;67(3):453-460.
38. Olivares JM, Alptekin K, Azorin JM, Canas F, Dubois V, Emsley R, Gorwood P, Haddad PM, Naber D, Papageorgiou G, Roca M, Thomas P, Martinez G, Schreiner A. Psychiatrists' awareness of adherence to antipsychotic medication in patients with schizophrenia: results from a survey conducted across Europe, the Middle East, and Africa. *Patient Preferences and Adherence* 2013;7:121-132.
39. Lambert M, Conus P, Eide P, Mass R, Karow A, Moritz S, Golks D, Naber D. Impact of present and past antipsychotic side effects on attitude toward typical antipsychotic treatment and adherence. *European Psychiatry* 2004;19(7):415-422.
40. Perkins DO, Gu H, Weiden PJ, McEvoy JP, Hamer RM, Lieberman JA. Predictors of treatment discontinuation and medication nonadherence in patients recovering from a first episode of schizophrenia, schizophreniform disorder, or schizoaffective disorder: a randomized, double-blind, flexible-dose, multicenter study. *Journal of Clinical Psychiatry* 2008;69(1):106-113.
41. Fervaha G, Takeuchi H, Agid O, Lee J, Foussias G, Remington G. Determinants of patient-rated and clinician-rated illness severity in schizophrenia. *Journal of Clinical Psychiatry* 2015;76(7):924-930.
42. Fervaha G, Agid O, Takeuchi H, Lee J, Foussias G, Remington G. Relationship between symptomatic improvement and overall illness severity in patients with schizophrenia. *Journal of Clinical Psychopharmacology* 2015;35(2):128-133.

43. Vanelli M, Burstein P, Cramer J. Refill patterns of atypical and conventional antipsychotic medications at a national retail pharmacy chain. *Psychiatric Services* 2001;52(9):1248-1250.
44. Perkins DO. Predictors of noncompliance in patients with schizophrenia. *Journal of Clinical Psychiatry* 2002;63(12):1121-1128.

Table 6.1

Descriptive statistics and inter-correlations of baseline characteristics in the study sample

Baseline Characteristics	Mean	SD	N	Pearson correlation coefficient [i]							
				Education	Marital status	Substanceab use	Diagnosis	Years symptoms	Hospital. at BL	CGI-S at BL	Treatment drug
Gender [a]	0.54	0.50	1136	0.02	0.17**	0.16**	-0.13**	-0.14**	-0.03	0.06*	-0.05
Education [b]	10.22	2.21	1136		0.02	0.04	-0.04	-0.09**	0.05*	-0.02	0.01
Marital status [c]	0.79	0.41	1136			0.09**	-0.17**	-0.12**	0.10**	0.05*	-0.01
Substance abuse [d]	1.08	0.34	1136				0.01	-0.08**	0.07**	0.11**	0.04
Diagnosis [e]	1.14	0.35	1136					0.03	-0.09**	0.02	0.01
Years symptoms	10.70	9.24	1134						-0.04	0.04	0.05
Hospitalized at BL [f]	0.46	0.50	1136							0.20**	0.02
Severity CGI-S at BL [g]	4.80	0.69	1136								0.01
Treatment drug [h]	0.75	0.43	1136								

a 0=Female (n=528), 1=Male (n=608);

b Number of years;

c 0=Married, or with partner (n=238), 1=Single (n=898);

d 1=Abstinence (n=1062), 2=Use without impairment (n=54), 3=Abuse (n=19),4=Dependence (n=1);

e 1=Schizophrenia (n=977), 2=Schizoafffective disorder (n=159);

f BL=Baseline, 0=Not hospitalized (n=613), 1=Hospitalized (n=523);

g CGI-S=Clinical Global Impression-Severity scale: 3=Mildly ill (n=1), 4=Moderately ill (n=396), 5=Markedly ill (n=567), 6=Severely ill (n=169), 7=Among most severely ill (n=3);

h 0=Olanzapine (n=281), 1=Asenapine (n=855);

i * p<0.05, ** p<0.01.

Table 6.2

Descriptive statistics and intercorrelations of efficacy and safety parameters

Effectiveness / Treatment engagement	Mean	SD	N	Pearson correlation coefficient [i]								
				Global Improvement	Insight	Positive factor	Negative factor	Disorganized factor	Excited factor	Emotional factor	SWN total	SF-12 Physical	SF-12 Mental
Severe AEs [a]	0.10	0.44	1136	0.13**	0.05*	0.13**	0.07*	0.07**	0.11**	0.19**	-0.05*	-0.06*	-0.05
Global Improvement [b]	2.89	1.63	1128		0.43**	0.59**	0.49**	0.55**	0.54**	0.47**	-0.32**	-0.21**	-0.19**
PANNS item/factor score [c]													
Insight [d]	2.97	1.30	1103			0.65**	0.58**	0.77**	0.62**	0.32**	-0.14**	-0.11**	-0.08**
Positive factor [e]	12.04	5.71	1104				0.59**	0.75**	0.74**	0.57**	-0.28**	-0.21**	-0.20**
Negative factor [f]	21.03	7.51	1104					0.80**	0.50**	0.58**	-0.32**	-0.18**	-0.24
Disorganized factor [g]	16.84	6.03	1104						0.72**	0.54**	-0.22**	-0.17**	-0.16**
Excited factor [h]	7.97	4.00	1104							0.49**	-0.20**	-0.11**	-0.13**
Emotional factor [i]	8.45	3.91	1104								-0.41**	-0.23**	-0.37**
SWN total	83.56	19.00	981									0.40**	0.46**
SF-12													
Physical	45.42	8.67	967										0.23**
Mental	42.50	9.59	967										

a Occurrence (number) of severe adverse events, at least possibly related to treatment: 0 in 1058 participants (pts), 1 in 59 pts, 2 in 10 pts, 3 in 5 pts, 4 in 3 pts, 6 in 1 pts;

b CGI-I at endpoint (EP): 1=Very much improved (n=252), 2=Much improved (n=327), 3=Minimally improved (n=184), 4=No change (n=144), 5=Minimally worse (n=109), 6=Much worse (n=99), 7=Very much worse (n=13), values missing n=8;

c Positive and Negative Syndrome Scale (PANSS) scores at EP;

d PANSS item 'lack of judgment and insight' (range 1-7);

e Subtotal of PANSS items 'delusions', 'hallucinations', 'grandiosity', 'suspiciousness', 'unusual thought content' (range 5-35);

f Subtotal of PANSS items 'conceptual disorganization', 'blunted affect', 'emotional withdrawal', 'poor rapport', 'social apathy', 'lack of spontaneity', 'motor retardation', 'active social avoidance' (range 8-56);

g Subtotal of PANSS items 'conceptual disorganization', 'stereotyped thinking', 'mannerism and posturing', 'disorientation', 'poor attention', 'lack of insight', 'preoccupation' (range 7-49);

h Subtotal of PANSS items 'excitement', 'hostility', 'uncooperativeness', 'poor impulse control' (range 4-28); i Subtotal of PANSS items 'anxiety', 'guilt', 'tension', 'depression' (range 4-28); j * p<0.05, ** p<0.01.

Table 6.3*a*

Joint effect of baseline characteristics, outcome, and satisfaction with treatment on study discontinuation (ITE, WOC)

Dependent variable	Model [a]	Significant predictors [b]	Stand. β [c]	p	Model information			R square	
					Test value	d.f.	p	Change	Total
Discontinuation: ITE	I: Risk factors				F = 8.07	9	<0.001	0.070	0.061
		CGI - Baseline severity ↑	0.065	0.013					
	II: I + Effectiveness				F = 49.12	12	<0.001	0.309	0.371
		CGI-Improvement ↓	0.083	0.013					
		IST ↓	-0.598	<0.001					
	III: II + Treatment Engagement				F = 44.43	14	<0.001	0.013	0.383
		Insight (PANSS item G12) ↓	0.082	0.007					
		PST ↑	0.144	<0.001					
Discontinuation: WOC	I: Risk factors				F = 1.24	9	0.270	0.011	0.002
		CGI - Baseline severity ↓	-0.071	0.025					
	II: I + Effectiveness				F = 2.09	12	0.015	0.014	0.013
		CGI-Improvement –	-0.011	0.776					
		IST ↑	0.260	<0.001					
	III: II + Treatment engagement				F = 9.80	14	<0.001	0.099	0.112
		Insight (PANSS item G12) –	0.054	0.139					
		PST ↓	-0.478	<0.001					

ITE Insufficient therapeutic effect; WOC Withdrawal of consent; IST Investigator's satisfaction with treatment; PST Patient's satisfaction with treatment; CGI Clinical Global Impression; PANSS Positive and Negative Syndrome Scale.

a Subsequent entering of blocks I-IV in the models for ITE and WOC was associated with a R square total of less than 0.07 and not listed here. The + sign indicates that the variables listed are added to those in previous blocks.

b Only independent variables attaining a significant standardized β value in the full model for ITE or WOC are listed here. The direction of the effect is indicated by arrows (↑ positive effect, ↓ negative effect, – no significant effect on dependent variable).

c Standardized β values in the final model (VI).

Table 6.3*b*

Joint effect of baseline characteristics, outcome, and satisfaction with treatment on satisfaction with treatment

Dependent variable	Model[a]	Significant predictors[b]	Stand. B[c]	p	Model information			R square	
					Test value	d.f.	p	Change	Total
Satisfaction: IST	I: Risk factors				F = 13.04	9	<0.001	0.078	0.068
		CGI - Baseline severity ↑	0.085	<0.001					
		Schizoaffective disorder ↓	-0.073	0.002					
	II: I + Illness (clinician-observed)				F = 58.70	15	<0.001	0.504	0.574
		PANSS positive factor ↓	-0.293	<0.001					
		PANSS negative factor ↓	-0.128	0.001					
		PANSS disorganized factor ↓	-0.107	0.032					
		PANSS excited factor ↓	-0.201	<0.001					
		PANSS emotional factor ↓	-0.138	<0.001					
		Side effects (severe) ↓	-0.078	0.001					
	III: II + Impairment (self-rated)				F = 49.08	18	<0.001	0.002	0.575
		SF-12 Mental component –	0.027	0.296					
		SF-12 Physical component –	0.035	0.147					
Satisfaction: PST	I: Risk factors				F = 4.39	9	<0.001	0.043	0.033
		CGI - Baseline severity ↑	0.072	0.008					
		Schizoaffective disorder ↓	-0.061	0.022					
	II: I + Illness (clinician-observed)				F = 44.51	15	<0.001	0.391	0.424
		PANSS positive factor ↓	-0.155	0.001					
		PANSS negative factor ↓	-0.118	0.011					
		PANSS disorganized factor –	-0.070	0.224					
		PANSS excited factor ↓	-0.132	0.002					
		PANSS emotional factor ↓	-0.267	<0.001					
		Side effects (severe) –	-0.040	0.127					
	III: II + Impairment (self-rated)				F = 38.74	18	<0.001	0.012	0.434
		SF-12 Mental component ↑	0.064	0.031					
		SF-12 Physical component ↑	0.062	0.029					

IST Investigator's satisfaction with treatment; PST Patient's satisfaction with treatment; CGI Clinical Global Impression; PANSS Positive and Negative Syndrome Scale; SF-12 Short Form health survey (12-item).

a Subsequent entering of blocks I-IV in the models for IST and PST was associated with a R square total of less than 0.07 and not listed here. The + sign indicates that the variables listed are added to those in previous blocks.

b Only independent variables attaining a significant standardized β value in the full model for IST or PST are listed here. The direction of the effect is indicated by arrows (↑ positive effect, ↓ negative effect, – no significant effect on dependent variable).

c Standardized β values in the final model (VI).

Chapter 7

Comparison of drugs across two ethno-geographical regions in depression

Demonstration of equal efficacy & safety as to avoid repeat studies [1]

[1] This chapter was published as:

Murasaki, M. *(h.c.)*, Schoemaker, J., Miyake, M.., Gailledreau, J., Heukels, A.J., Fennema, H.P., Sitsen, J.M.A. (2010).

Comparison of efficacy and safety of mirtazapine versus fluvoxamine in Japanese and caucasian patients with major depressive disorder. *Japanese Journal of Clinical Psychopharmacology, 13,*339-355

ABSTRACT

Introduction: Primary aim was to demonstrate non-inferiority of mirtazapine compared to fluvoxamine, using a 2 points margin of difference in summary score change from baseline on the first 17 items of Hamilton Depression Rating Scale ($HAMD_{17}$) in- and outside Japan. Secondary objective was to compare the safety profile of both mirtazapine in Japanese (J) and Caucasian (C) patients.

Methods: A randomized, double-blind, 6-week, multi-center trial was conducted with flexible doses of mirtazapine (15-45 mg/day) and fluvoxamine (50-150 mg/day) in patients suffering from major depressive disorder in Japan and Europe.

Results: In total 402 patients (20-73 years) were randomized and treated, of which 199 on mirtazapine (J=96, C=103), and 203 on fluvoxamine (J=98, C=105). A clinically significant decrease in mean $HAMD_{17}$ total score was observed on mirtazapine (-14.3) and fluvoxamine (-13.6). The observed -0.7 points difference in $HAMD_{17}$ total score change from baseline in favor of mirtazapine was not statistically significant, but allowed a conclusion of non-inferiority. There was a tendency for Japanese patients to respond slightly better to mirtazapine, and for Caucasian patients to respond slightly better to fluvoxamine, but the herewith presumed treatment*region interaction was not statistically significant (P=0.075) and a possible result of random variation. Mirtazapine was characterized with a faster onset of action as compared to fluvoxamine. Both drugs were generally well tolerated. Mirtazapine was associated with a higher frequency of somnolence in Japanese patients, whereas fluvoxamine was associated with a higher incidence of nausea in both Japanese and Caucasian patients.

Conclusion: Mirtazapine was shown to be non-inferior to fluvoxamine in the treatment of major depressive disorder. Both drugs were well tolerated with some adverse events being almost exclusively reported in Japanese patients and some in Caucasian patients.

INTRODUCTION

Because of improved tolerability as compared to tricyclic antidepressants (TCA), selective serotonin reuptake inhibitors (SSRIs) have taken their place in the first line treatment of major depressive disorder.[1-3] One of the first SSRIs entering the market was fluvoxamine, available in many countries throughout the world and available in Japan since 1999. Although fluvoxamine is an effective antidepressant, its clinical use is associated with frequent occurrence of gastro-intestinal disorders (e.g. nausea) and the requirement of careful administration because of drug interaction caused by its cytochrome P450 (CYP1A2, 2D6, and 3A) inhibitory activity.[4-6]

Mirtazapine is a novel antidepressant with specific antagonistic activity against central α_2-autoreceptor and 5-HT_2 and 5-HT_3 receptors.[7,8] The blockade of presynaptic α_2-autoreceptor causes an increased release of noradrenalin. The subsequent excitation of postsynaptic α_1-adrenoceptors mediates 5-HT cell firing and the direct blockade of inhibitory α_2-heteroreceptors located on 5-HT terminals facilitate serotonergic neurotransmission. The effect of the released serotonin is exerted only via 5-HT_{1A} receptors, since 5-HT_2 and 5-HT_3 receptors are blocked by mirtazapine. Since mirtazapine has low Ki for CYP1A2, 2D6, and 3A (Ki: 41-210 μM), drug interactions observed with fluvoxamine are not expected.[9-11]

In the clinical development program the antidepressant activity of mirtazapine had been demonstrated to be superior to that of placebo, comparable to that of amitriptyline and other currently available tricyclic antidepressants (TCAs), and superior to that of trazodone.[12-20] In addition, non-inferiority to SSRIs (sertraline, paroxetine, citalopram and fluoxetine) had been demonstrated.[15,21-23] So far, none of these trials with mirtazapine had been performed in Japanese patients, so that additional studies were deemed required before mirtazapine could be allowed to the market as antidepressant in Japan.

In order expand the registration of mirtazapine to Japan, an active-controlled study was considered useful as to demonstrate the effectiveness of mirtazapine in the Japanese population, while at the same time offering the possibility for adequate treatment to all trial participants. Since at the time of study design fluvoxamine was the only SSRI approved for first line treatment for depression in Japan, fluvoxamine was chosen as active comparator in a non-inferiority setting. Centers in Europe were included to allow comparison of results obtained in Japanese and Caucasian patients treated under the same protocol. If it could be

unequivocally demonstrated that mirtazapine is not inferior to fluvoxamine in both populations, and equally well tolerated in Japanese as it is in Caucasians, results from the full clinical development program of mirtazapine in Caucasians might be considered equally relevant for the treatment of depression in Japan, and in support of its allowance in this region. In other words, it was anticipated that on the basis of a single bridging study, extrapolation of clinical data obtained in foreign countries was possible and could satisfy the requirements for pre-registration evidence in Japan, without conducting one or more supplementary (placebo-controlled) studies.

Here we report on the results from a non-inferiority, randomized, double-blind, multi-center trial, comparing the outcome of six weeks, flexible-dose treatment with mirtazapine (15-45 mg/day) and fluvoxamine (50-150 mg/day) in depressed outpatients across Japan and Europe. Efficacy was primarily compared using the summary score on the first 17 items of the Hamilton Depression Rating Scale ($HAMD_{17}$) by end of treatment (last observation carried forward, or LOCF), with a one-sided non-inferiority margin of maximally 2.0 points.

METHODS

Study population

The objectives of the study were (*1*) to compare the efficacy of mirtazapine and fluvoxamine in the treatment of major depressive disorder, and (*2*) to compare the safety profile in terms of adverse event (AE) patterns on both drugs in Japanese and Caucasian patients.

A sample of 400 patients (200 patients for each treatment group) was considered adequate to establish non-inferiority of mirtazapine. One hundred patients for each group (200 in total) were to be recruited in Europe; the other half of the patients (100 patients for each group, 200 in total) were to be recruited in Japan.

Patients were eligible for enrolment when they (*1*) were at least 20 years but not older than 75 years of age at the time of screening; (*2*) fulfilled the DSM-IV criteria for a Major Depressive Episode (DSM-IV diagnosis: 296.2x or 296.3x); (*3*) showed a minimum score of 18 on HAM-D_{17} at screening and baseline; (*4*) had the present episode of depression for at

least two weeks but no longer than 12 months; (*5*) were willing to provide written informed consent.

Patients were excluded from study participation when they (*1*) suffered from any medical condition or required concomitant treatment that would confound study results or put the patient at increased risk of treatment failure or unacceptable adverse events; (*2*) were hospitalized or required hospitalization at screening or baseline; (*3*) were pregnant or lactating; (*4*) were women of childbearing potential while not practicing a reliable method of contraception; (*5*) had a score of 3 or more on the HAMD item [11] 'suicide'; (*6*) had a history or present condition of bipolar disorder, schizophrenia or psychotic symptoms, schizotypal or borderline personality disorder, or organic mental disorder; (*7*) had a present condition of anxiety disorder (according to DSM-IV), eating disorder, postpartum depression, epilepsy or history of seizure disorder, or ever received treatment with anticonvulsant medication for epilepsy or seizures, alcohol or substance abuse (according to DSM-IV) during the last six months; (*8*) had any clinically meaningful non-stable renal, hepatic, cardiovascular respiratory, cerebrovascular disease or other serious, progressive physical disease; (*9*) had any clinically meaningful abnormal finding uncovered during the physical examination, ECG and/or clinically significant abnormal laboratory results at screening; (*10*) had participated in other clinical trials within the last three months; (*11*) had received any of the following treatments (within the indicated intervals between brackets) before the start of active treatment: non selective MAO inhibitors or lithium (three weeks), depot neuroleptics (three months), other psychotropics (two weeks), ECT (three months), or other investigational drugs (three months); (*12*) had a body mass index greater than 30 or less than 18; (*13*) had a known allergy or hypersensitivity to mirtazapine or fluvoxamine; (*14*) had been treated during the present episode with either mirtazapine or fluvoxamine; (*15*) required treatment with terfenadine, astemizole, warfarin, theophylline, or cisaprid; (*16*) received concurrent treatment for a micturition disorder.

The study was conducted in compliance with the Declaration of Helsinki, Good Clinical Practice (GCP), and other ordinance. Before the start of screening, written informed consent was obtained from each participant after the investigator, or another appropriate individual who was designated by the investigator, had informed the subject about the nature and purposes of the study, as well as the trial procedures in both written and oral form.

Study design

The study was a randomized, double-blind, multicenter, flexible dose, parallel group study. Randomization to treatment was stratified within center. Patients fulfilling all inclusion criteria and violating none of the exclusion criteria were randomly allocated to six weeks treatment with mirtazapine once daily, or to six weeks treatment with fluvoxamine twice daily. Mirtazapine was to be dosed in the evening (at approximately 21:00 hours). Fluvoxamine was to be dosed in the morning after breakfast and in the evening (at approximately 21:00 hours). The study drugs were administered using a double-dummy technique to guarantee blinding to treatment. During the first week of treatment, patients allocated to mirtazapine were to receive a daily dose of 15 mg. Patients allocated to fluvoxamine were to receive 25 mg in the morning and 25 mg in the evening. If, after one week of treatment, there was insufficient response to the medication (as defined by a decrease of less than 20% on the HAM-D_{17} total score from baseline) and tolerability was acceptable, the investigator was allowed to increase the daily dose to either 30 mg mirtazapine or 100 mg fluvoxamine. From the third week onwards, provided tolerability remained acceptable and response was still not satisfactory in the opinion of the investigator, the dosage could be further increased to either 45 mg mirtazapine or 150 mg fluvoxamine. In case of tolerability problems, the dosage could be reduced again. However, a lower dose than the minimal starting dose was not allowed. The total daily dose, therefore, could vary between 15 mg and 45 mg for mirtazapine, and between 50 mg and 150 mg for fluvoxamine. Clinical assessments were carried out at the screening visit, the baseline visit, and at weekly intervals during the study (days 7, 14, 21, 28, 35, and 42 with an allowance for each visit after baseline of ±2 days).

Efficacy

The difference in change from baseline on $HAMD_{17}$ by treatment was used as the primary outcome variable. The rating scale actually used in this trial was the semi-structured, anchored, 21-item version as described by Williams.[24] All participating investigators attended a training session to standardize its use. Efficacy was also assessed using the Clinical Global Impression scale (CGI), which consists of a seven-stage rating of the severity of the illness within the preceding two days (CGI-S), and a global seven-stage ratings of the change in clinical status since the start of treatment (CGI-I).[25] This scale was completed at baseline

(only for severity), and at all further assessments during the treatment period (days 7, 14, 21, 28, 35, and 42 with an allowance for each visit after baseline of ±2 days).

Safety

ECG and physical examinations were evaluated at screening and at day 42 (or as soon as possible after premature discontinuation). Routine laboratory tests were performed at screening, day 14 and day 42 (or as soon as possible after premature discontinuation). Information regarding AEs was to be obtained by questioning or examining the patient at each visit during the treatment period. AEs were also collected from spontaneous reporting.

Statistical analysis

The primary objective was to establish non-inferiority of mirtazapine versus fluvoxamine at endpoint (EP) of treatment on the change from baseline on $HAMD_{17}$ summary score. Mirtazapine was considered non-inferior if the upper limit of the 95% CI of the difference between the two treatments was less than 2.0 points. Only patients who were randomized and treated with active medication, and who had a baseline measurement and at least one post-baseline efficacy measurement were included in the efficacy analysis (Intent-To-Treat, or ITT group). The estimate of treatment effect and corresponding (two-sided) 95% confidence intervals was based on the additive two-way ANOVA with factors for treatment and region. Region was defined as the geographical region of centers in Japan (Hokkaido, five centers; Saitama, four centers; Tokyo, seven centers; Kanagawa, nine centers; Shin-Etsu, three centers, Western Japan, six centers) and countries in Europe (Denmark, four centers; France, three centers; Belgium and the Netherlands, six centers; Norway, four centers; Spain, three centers). Note that by adding region the analysis was automatically stratified for ethnicity. A test for treatment by region interaction was planned by extending the model with the interaction term. In case of a significant interaction ($p \leq 0.10$) the differences between the regions with respect to the treatment effects were to be presented, discussed and further explored to evaluate whether it was still justified to present an overall estimate of the treatment effect. To investigate treatment differences between Japanese and Caucasian patients, an additive two-way ANOVA with a factor for treatment, ethnic group (Japanese and Caucasian) and treatment by ethnic group interaction was planned as an exploratory analysis.

Response was defined by least 50% reduction in $HAMD_{17}$ score from baseline to EP, and a score of <3 ('much' or 'very much improved') on CGI-I.

For comparison of safety within (by region) and between treatment groups, adverse event frequencies and observed laboratory values outside the reference range are listed for all subjects who were randomized and treated with active medication (All-Subjects-Treated, or AST group).

In order to explore heterogeneity between treatment groups, demographics and baseline characteristics were tabulated by drug and overall (pooled across ethnic groups). In addition, analyses of covariate adjustment for baseline factors with treatment imbalances at baseline (i.e. gender and disease condition) were performed for CGI-S, $HAMD_{17}$, and HAMD Factor VI 'sleep disturbance' scores.

For the evaluation of safety, descriptive statistics in the ITT group were calculated for adverse events, laboratory parameters, physical examinations, vital signs and ECG results. No inferential tests were performed.

All analyses were done using SAS Version 6.12 under Windows on a PC system.

RESULTS

Study population

From a total of 434 patients who were screened for eligibility, 412 patients were randomized to treatment between August 2000 and March 2002. Because of GCP violation at one site in Japan, nine of these patients (five on mirtazapine, and four on fluvoxamine) were excluded from further analysis. From the remaining 403 randomized patients, one patient on mirtazapine dropped out before having taken any trial medication, leaving 402 patients in the all-patients-treated (AST) group. In five patients efficacy was not assessed after baseline, so that 397 patients remained available in the ITT group for efficacy analysis. Overall, both treatment groups were comparable with respect to number of allocated patients (Figure 7.1).

The average baseline characteristics by treatment and region are provided in Tables 7.1*a* and 7.1*b*. Most remarkable differences were that in Japan, patients had a lower body mass index, appeared less severely depressed (as reflected by baseline $HAMD_{17}$ and CGI-S scores),

and were more frequently of male gender compared to the patients in Europe. Age and duration of the present episode of depression was similar between and within treatment groups.

Exposure was comparable between and within treatments groups (Table 7.2). Mean duration of exposure was 36.7 days on mirtazapine (Japanese: 34.9 days, Caucasian: 38.4 days) and 37.5 days on fluvoxamine (Japanese: 35.9 days, Caucasian: 39.0 days). The mean daily dose (SD), including the 1st-week fixed-dose period was 21.0 (7.7) mg for mirtazapine (Japanese: 19.2 (8.5) mg, Caucasian: 22.7 (6.5) mg) and 78.5 (26.3) mg for fluvoxamine (Japanese: 78.9 (29.8) mg, Caucasian: 78.1 (22.6) mg) (Table 7.2). There were no major differences in dosing between the two ethnic groups. Overall compliance with the intake of study medication was good with no important differences between the ethnic and treatment groups.

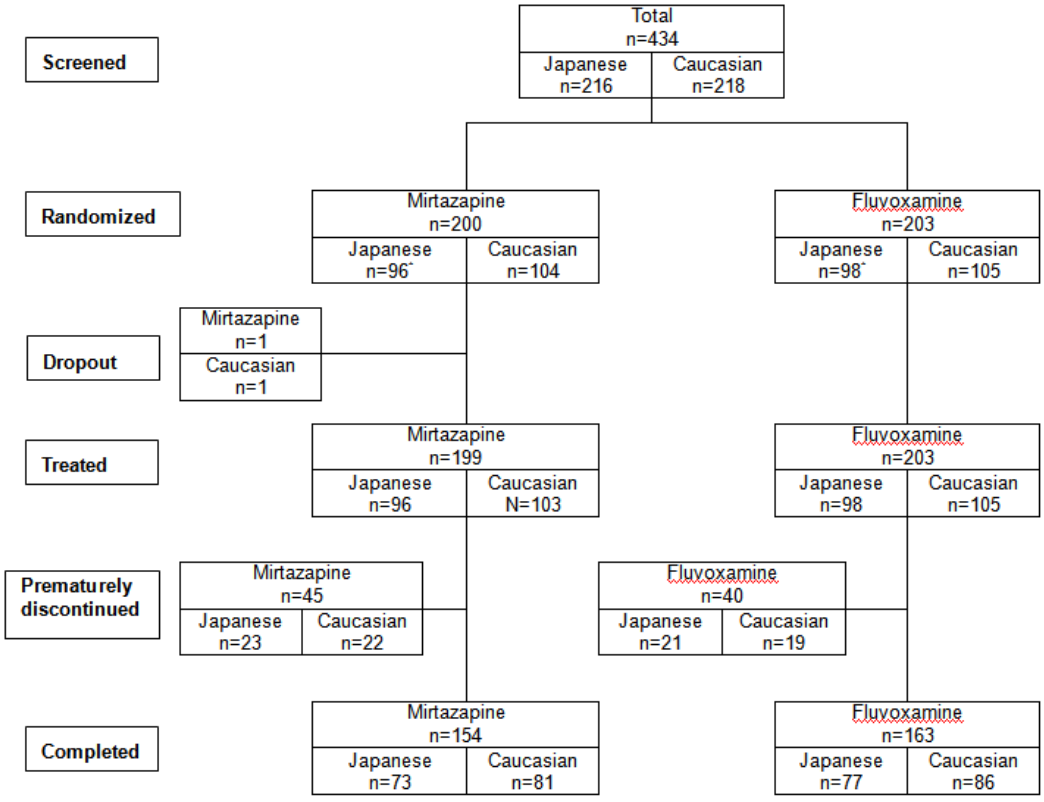


Figure 8.1

Disposition of subjects

* Because of GCP violation at one site in Japan, 12 patients enrolled at that site were excluded.

The percentage of patients who prematurely discontinued the study was 22.7% for Japanese (mirtazapine: 24.0%, fluvoxamine: 21.4%) and 19.7% for Caucasians (mirtazapine: 21.4%, fluvoxamine: 18.1%). Due to premature discontinuations, the percentage of patients receiving the drug for 10 days or less in Japanese patients (mirtazapine: 15.6%, fluvoxamine: 10.2%) was higher than that in Caucasian patients (mirtazapine: 6.8%, fluvoxamine: 5.7%).

Table 7.1*a*

Study population: baseline characteristics (AST group)

	Caucasian		Japanese		
AST group	**Mirtazapine (n=103)**	**Fluvoxamine (n=105)**	**Mirtazapine (n=96)**	**Fluvoxamine (n=98)**	**Total (N=402)**
Gender					
Female	70 (68.0%)	67 (63.8%)	49 (51.0%)	38 (38.8%)	224 (55.7%)
Male	33 (32.0%)	38 (36.2%)	47 (49.0%)	60 (61.2%)	178 (44.3%)
Age (years					
Mean (SD)	42.0 (12.1)	39.4 (12.7)	41.3 (13.2)	42.2 (12.8)	41.2 (12.7)
Range	20 - 69	20 - 71	20 - 72	20 - 73	20 - 73
Height (cm)					
Mean (SD)	169.5 (9.6)	169.7 (8.2)	163.2 (8.2)	164.2 (9.7)	166.8 (9.4)
Range	147 - 192	153 - 190	146 - 182	142 - 186	142 - 192
Body weight (kg)					
Mean (SD)	68.6 (12.1)	68.4 (12.4)	59.5 (9.5)	60.9 (11.0)	64.5 (12.1)
Range	45.0 - 106.0	44.0 - 102.0	40.2 - 88.0	38.0 - 89.8	38.0 - 106.0
Body mass index [kg/m^2]					
Mean (SD)	23.7 (3.1)	23.7 (3.4)	22.3 (2.7)	22.5 (2.9)	23.0 (3.1)
Median (range)	23.3 (18 - 30)	23.3 (17 - 33)	22.0 (18 - 30)	22.4 (17 - 29)	22.6 (17 – 33)
BMI by region	*Caucasian (n=205)*		*Japanese (n=194)*		
Mean (SD)	23.7 (3.3)		22.4 (2.8)		
Median (range)	23.3 (17 - 33)		22.2 (17 - 30)		
DSM-IV diagnosis					
296.2 (single episode)	44 (42.7%)	41 (39.0%)	33 (34.4%)	46 (46.9%)	164 (40.8%)
296.3 (recurrent)	59 (57.3%)	64 (61.0%)	63 (65.6%)	52 (53.1%)	238 (59.2%)
Duration present episode					
2 weeks - 1 month	12 (11.7%)	11 (10.5%)	10 (10.4%)	9 (9.2%)	42 (10.4%)
1 - 6 months	56 (54.4%)	62 (59.0%)	57 (59.4%)	53 (54.1%)	228 (56.7%)
6 - 12 months	35 (34.0%)	32 (30.5%)	29 (30.2%)	36 (36.7%)	132 (32.8%)
>12 months	0	0	0	0	0

Table 7.1*b*

Study population: baseline characteristics (ITT group)

ITT group	Mirtazapine (n=101)	Fluvoxamine (n=104)	Mirtazapine (n=95)	Fluvoxamine (n=98)	Total (n=398)
$HAMD_{17}$					
Mean (SD)	25.1 (3.67)	25.1 (4.09)	22.7 (3.36)	22.6 (3.90)	23.9 (3.95)
Range	19 - 38	18 - 40	18 - 31	16 - 34	16 - 40
Distribution					
< 18	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.0%)	1 (0.3%)
≥18, <25	50 (49.5%)	49 (47.1%)	72 (75.8%)	72 (73.5%)	243 (61.1%)
≥25	51 (50.5%)	55 (52.9%)	23 (24.2%)	25 (25.5%)	154 (38.7%)
$HAMD_{17}$ by region	*Caucasian (n=205)*		*Japanese (n=193)*		
Mean (SD)	25.0 (3.75)		22.6 (3.63)		
Range	18 - 40		16 - 34		
Distribution					
< 18	0 (0.0%)		1 (0.5%)		
≥18, <25	99 (48.3%)		144 (74.6%)		
≥25	106 (51.7%)		48 (24.9%)		
CGI-S					
mildly ill	2 (2.0%)	3 (2.9%)	14 (14.7%)	16 (16.3%)	35 (8.8%)
moderately ill	31 (30.7%)	38 (36.5%)	62 (65.3%)	66 (67.3%)	197 (49.5%)
markedly ill	58 (57.4%)	49 (47.1%)	17 (17.9%)	15 (15.3%)	139 (34.9%)
severely ill	10 (9.9%)	14 (13.5%)	2 (2.1%)	1 (1.0%)	27 (6.8%)

Table 7.2

Mean daily dose (mg) and extent of exposure (ITT group)

	Japanese		Caucasian	
	Mirtazapine (n=95)	Fluvoxamine (n=98)	Mirtazapine (n=101)	Fluvoxamine (n=104)
Daily dose (mg)				
Mean (SD)	19.2 (8.5)	78.9 (29.8)	22.7 (6.5)	78.1 (22.6)
Median (range)	15 (0 - 38)	88 (17 - 126)	24 (10 - 38)	82 (16 -139)
Daily dose by treatment	*Mirtazapine (n=196)*		*Fluvoxamine (n=202)*	
Mean (SD)	21.0 (7.7)		78.5 (26.3)	
Median (range)	22 (0 - 38)		83 (16 - 139)	
Exposure (days)				
Mean (SD)	34.9 (14.1)	35.9 (12.4)	38.4 (10.7)	39.0 (9.8)
Median (range)	42 (1 - 45)	42 (1 - 44)	42 (1 - 50)	43 (2 - 46)
Exposure by treatment	*Mirtazapine (n=196)*		*Fluvoxamine (n=202)*	
Mean (SD)	36.7 (12.5)		37.5 (11.2)	
Median (range)	42 (1 - 50)		42 (1 -46)	

The percentage of dropouts for drug-related reasons (i.e. adverse events or lack of efficacy) was 10.8% in Japanese vs. 14.9% in Caucasians (Table 7.3). Adverse drug experiences were cause for premature termination in 12.1% of mirtazapine-treated patients and 7.9% in fluvoxamine-treated patients. There were no relevant differences in dropout rates or reason for dropout between treatment groups.

Table 7.3

Premature discontinuations by reason

Single most important reason	Japanese						Caucasian					
	Mirtazapine (n=96)		Fluvoxamine (n=98)		Subtotal (n=194)		Mirtazapine (n=103)		Fluvoxamine (n=105)		Subtotal (n=208)	
	n	%	n	%	n	%	n	%	n	%	n	%
Adverse event	10	10.4	5	5.1	15	7.7	14	13.6	11	10.5	25	12.0
Study violation	1	1.0	1	1.0	2	1.0	0	0.0	0	0.0	0	0.0
Poor cooperation	2	2.1	2	2.0	4	2.1	6	5.8	3	2.9	9	4.3
Lack of efficacy	2	2.1	4	4.1	6	3.1	2	1.9	4	3.8	6	2.9
Other reason(s)	8	8.3	9	9.2	17	8.8	0	0.0	1	1.0	1	0.5
Total	23	24.0	21	21.4	44	22.7	22	21.4	19	18.1	41	19.7

Efficacy

In the pooled patient population, clinical improvement was evident in both treatment groups, with no statistically significant differences between drugs administered. The primary objective of non-inferiority of mirtazapine versus fluvoxamine at endpoint was met with an estimate of -0.69 (with a 95% confidence interval of -2.17 to 0.78) which was below the non-inferiority margin of 2.0 points. No statistically significant treatment by geographical region interaction was observed (P=0.406), suggesting that there was no significant difference in the treatment effect of both compounds across geographical regions (Table 7.4). The mean decrease in $HAMD_{17}$ total score on mirtazapine was similar in Japanese and Caucasian patients (-13. 8 and -14.4, respectively). On the other hand, the mean decrease in $HAMD_{17}$ total score on fluvoxamine was higher in Caucasian patients than in Japanese patients (-15.1 and -11.7, respectively). Japanese patients seemed to respond slightly better to mirtazapine, whereas Caucasian patients slightly better to fluvoxamine, but the herewith presumed treatment*region interaction was not statistically significant (P=0.075) and a possible result of random variation.

When comparing mirtazapine versus fluvoxamine on $HAMD_{17}$ total score change from baseline by time of assessment, a greater extent of reduction was observed in the mirtazapine

group than in the fluvoxamine group in the early stages of the trial, with statistically significant differences at day 7: -1.97 (P<0.001) and day 14: -1.86 (P=0.001) (Figure 7.2, Table 7.5). This pattern was similar in both the Japanese and Caucasian populations. The data suggest an earlier onset of action by mirtazapine.

The HAMD responder rates (≥50% reduction on $HAMD_{17}$ total scores) in the ITT population at EP were comparable for mirtazapine (64.8%) and fluvoxamine (61.4%) (Table 7.6). The responder rate in Japanese was higher on mirtazapine (69.5%), but lower on fluvoxamine (55.1%) compared to the responder rates in Caucasians (60.4% and 67.3 % respectively).

The CGI responder rates were similar for mirtazapine as compared to fluvoxamine with a similar tendency of early response as observed on $HAMD_{17}$ (Table 7.7). In the fluvoxamine group, there was an apparent difference in responder rates between Caucasians and Japanese with a more pronounced responder rate in Caucasians (72.1% and 55.1% at EP, respectively) which was not observed in the mirtazapine group (62.4% and 60.0% at EP, respectively).

Safety

Adverse events reported for at least 5% of the patients in either treatment group during the treatment period (+5 days) are shown by WHO preferred term including causal relationship to the study drug as assessed by the investigator in Table 7.8.

The most frequently reported adverse events (AEs) in patients treated with mirtazapine were somnolence (28%) (Japanese: 44%, Caucasians: 13%), followed by dizziness (12%) (Japanese: 9%, Caucasians: 14%), thirst (11%) (Japanese: 22%, Caucasians: 1%), headache (11%) (Japanese: 12%, Caucasians: 10%), nausea (8%) (Japanese: 6%, Caucasians: 10%), weight increase (8%) (Japanese: 6%, Caucasians: 10%), malaise 7% (Japanese: 15%, Caucasians: 0%), fatigue 7% (Japanese: 0%, Caucasians: 14%), and constipation 7% (Japanese: 13%, Caucasians: 2%).

AEs occurring in at least 5% of the subjects of a treatment group are listed in Table 7.8 by system-organ class and relationship to trial medication (in which “related” is judged by the investigator as definitely, probably, possibly or unlikely). The safety profile of fluvoxamine differed from mirtazapine with increased incidences of gastro-intestinal system disorders, in particular nausea, diarrhea, dyspepsia and abdominal pain. Also upper respiratory tract infections were more frequently reported on fluvoxamine. For mirtazapine the safety profile

included more pronounced incidences of somnolence malaise/fatigue, increased weight/appetite, and dry mouth/thirst.

The overall incidence of adverse events tended to be higher in Japanese than in Caucasians on both drugs. Notable differences in the distribution of drug-related adverse events between ethnic groups included a higher incidence (regardless of treatment) of somnolence, thirst, dyspepsia, and hypoaesthesia, but lower incidence of nausea, dizziness, diarrhea, and increased appetite in Japanese patients. None of the Caucasians reported drug-related malaise or constipation, while drug-related fatigue was reported by 12.6% of the Caucasian patients but by none of the Japanese patients.

For five patients serious adverse events were reported during the treatment period. For one Japanese patient a suicide attempt was reported while on fluvoxamine treatment. For European patients on mirtazapine two cases of aggravated depression and a suicide attempt and a case of vaginitis were reported; for one patient on fluvoxamine an aggressive reaction was reported.

Table 7.4

$HAMD_{17}$ total score change from baseline by treatment and region (ITT, LOCF)

		LS mean	SE	d	95% CI	P
By treatment						
	Mirtazapine	-14.26	0.75	-0.69	-2.17 - 0.78	0.406
	Fluvoxamine	-13.57				
By region						
Japan						
	Mirtazapine	-13.77	0.78			
	Fluvoxamine	-11.71	0.77	-2.73	-5.73 - 0.28	0.075
Europe						
	Mirtazapine	-14.39	0.76			
	Fluvoxamine	-15.06	0.75			
By treatment per region						
Japan						
	Mirtazapine	-13.77	1.113	-2.06	-4.25 - 0.14	0.066
	Fluvoxamine	-11.71				
Europe						
	Mirtazapine	-14.39	1.049	0.67	-1.40 - 2.74	0.523
	Fluvoxamine	-15.06				

LS = least square; SE = standard error; d = estimated difference; CI = confidence interval; P = probability of significance

Table 7.5

$HAMD_{17}$ total score change from baseline over time by treatment and region (ITT, LOCF)

	Day 7	Day 14	Day 21	Day 28	Day 35	Day 42
By treatment						
Mirtazapine	-5.7 ± 5.16	-9.1 ± 5.86	-11.0 ± 6.70	-12.0 ± 7.08	-13.4 ± 7.26	-14.2 ± 7.37
Fluvoxamine	-3.7 ± 4.51	-7.3 ± 6.13	-10.0 ± 6.49	-11.6 ± 7.20	-12.6 ± 7.60	-13.6 ± 7.82
d	-1.97*	-1.86*	-1.09	-0.38	-0.89	-0.71
(95% CI)	(-2.90, -1.03)	(-3.00, -0.72)	(-2.34, 0.16)	(-1.78, 1.01)	(-2.34, 0.55)	(-2.19, 0.76)
By region						
Japan						
Mirtazapine	-6.0 ± 4.89	-9.1 ± 5.71	-10.3 ± 6.40	-12.0 ± 6.76	-13.2 ± 7.03	-14.1 ± 7.14
Fluvoxamine	-3.4 ± 4.57	-6.8 ± 5.96	-8.9 ± 6.55	-10.4 ± 7.35	-11.1 ± 7.78	-12.0 ± 7.97
d	-2.60*	-2.38*	-1.51	-1.71t	-2.33*	-2.27*
(95% CI)	(-3.91, -1.29)	(-4.01, -0.75)	(-3.35, 0.33)	(-3.72, 0.31)	(-4.44, -0.22)	(-4.42, -0.12)
Europe						
Mirtazapine	-5.4 ± 5.40	-9.2 ± 6.03	-11.7 ± 6.93	-11.9 ± 7.40	-13.6 ± 7.49	-14.4 ± 7.61
Fluvoxamine	-3.9 ± 4.47	-7.8 ± 6.27	-10.9 ± 6.31	-12.7 ± 6.91	-14.0 ± 7.18	-15.1 ± 7.41
d	-1.38*	-1.38	-0.70	0.84	0.42	0.72
(95% CI)	(-2.71, -0.05)	(-2.98, 0.22)	(-2.42, 1.02)	(-1.09, 2.77)	(-1.56, 2.41)	(-1.30, 2.73)

d = estimated difference between treatments; CI = confidence interval
* = $P < 0.05$
t = trend: $0.05 < P < 0.10$

Table 7.6

Frequency distribution of $HAMD_{17}$ responders (ITT group, LOCF and EP approach)

	Day 7		Day 14		Day 21		Day 28		Day 35		Day 42		EP	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
By treatment														
Mirtazapine	21	10.8	63	32.5	99	51.0	101	52.1	120	61.9	127	65.5	127	64.8
Fluvoxamine	18	9.0	51	25.5	88	44.0	109	54.5	111	55.5	124	62.0	124	61.4
By region														
Japan														
Mirtazapine	12	12.9	34	36.6	50	53.8	54	58.1	60	64.5	66	71.0	66	69.5
Fluvoxamine	11	11.5	23	24.0	38	39.6	45	46.9	46	47.9	54	56.3	54	55.1
Europe														
Mirtazapine	9	8.9	29	28.7	49	48.5	47	46.5	60	59.4	61	60.4	61	60.4
Fluvoxamine	7	6.7	28	26.9	50	48.1	64	61.5	65	62.5	70	67.3	70	67.3

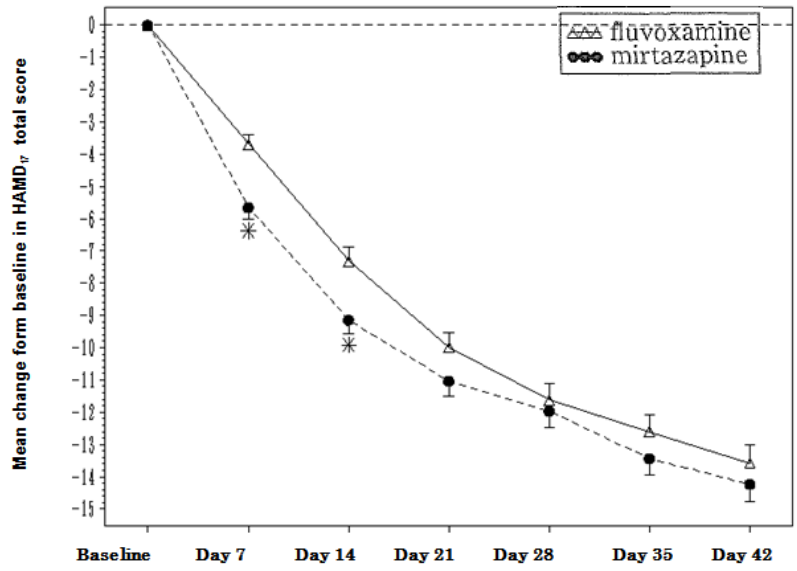


Figure 7.2

Mean change ± SE from baseline $HAMD_{17}$ by treatment over time (ITT group)

* P < 0.05 (ANOVA)

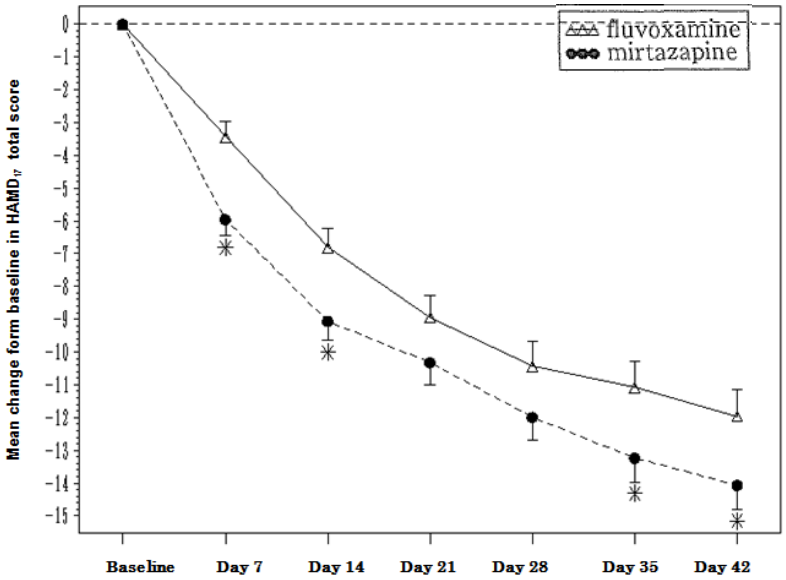


Figure 7.3

Mean change ± SE from baseline $HAMD_{17}$ by treatment and time (ITT group) in Japan

* P < 0.05 (ANOVA)

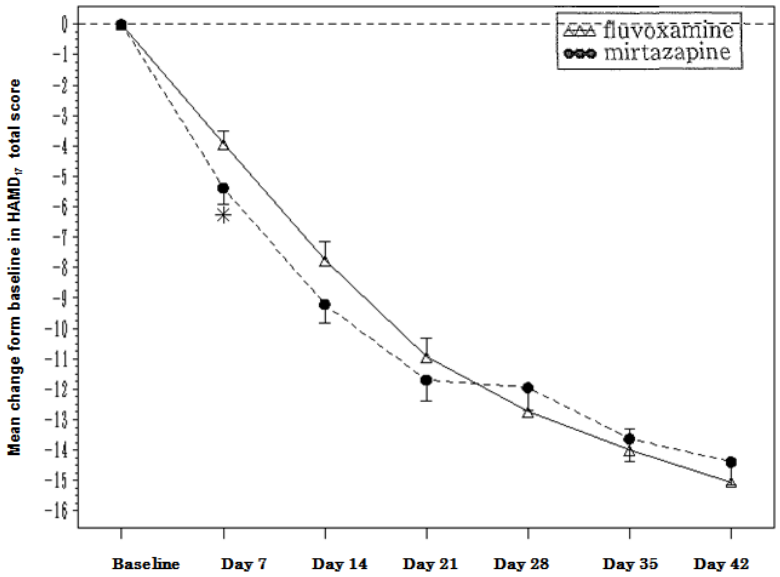


Figure 7.4

Mean change ± SE from baseline $HAMD_{17}$ by treatment and time (ITT group) in Europe

* P < 0.05 (ANOVA)

Table 7.7

Frequency distribution of CGI responders (ITT group, LOCF and EP approach)

	Day 7		Day 14		Day 21		Day 28		Day 35		Day 42		EP	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
By treatment														
Mirtazapine	34	17.5	71	36.6	100	51.5	102	52.6	118	60.8	120	61.9	120	61.2
Fluvoxamine	22	11.0	59	29.5	83	41.5	106	53.0	118	59.0	129	64.5	129	63.9
By region														
Japan														
Mirtazapine	13	14.0	31	33.3	43	46.2	46	49.5	52	55.9	57	61.3	57	60.0
Fluvoxamine	7	7.3	25	26.0	31	32.3	41	42.7	45	46.9	54	56.3	54	55.1
Europe														
Mirtazapine	21	20.8	40	39.6	57	56.4	56	55.4	66	65.3	63	62.4	63	62.4
Fluvoxamine	15	14.4	34	32.7	52	50.0	65	62.5	73	70.2	75	72.1	75	72.1

Within the fluvoxamine treatment group, discontinuations were predominantly due to AEs of the category psychiatric disorders (e.g. mood changes) or gastro-intestinal system disorders (e.g. nausea). Discontinuations on mirtazapine were predominantly due to AEs of the category psychiatric disorders (e.g. sleepiness) or central and peripheral nervous system disorders (e.g. dizziness). Neither of the two antidepressants appeared to cause clinically relevant changes in results of physical examinations, vital signs (heart rate and blood pressure) and ECG

evaluations. No clinically significant abnormal values were observed for biochemistry parameters, haematocrit, hemoglobin, or platelet counts in either of the treatment groups. Plasma cholesterol values above 1.2 times the upper normal limit were observed in 12% of mirtazapine treated, and in 10% of fluvoxamine treated patients at endpoint, but increased plasma cholesterol levels were in most cases already present before treatment.

Table 7.8

Adverse events occurring in ≥5% of patients in at least one treatment group (in %)

	Fluvoxamine						Mirtazapine					
	Japanese (n=98)		Caucasian (n=105)		Total (n=203)		Japanese (n=96)		Caucasian (n=103)		Total (n=199)	
	ADR[1)]	AE	ADR[1)]	AE	ADR[1)]	AE	ADR[1)]	AE	ADR[1)]	AE	ADR[1)]	AE
Patients with ≥1 AE	68.4	77.6	61.9	65.7	65.0	71.4	78.1	80.2	51.5	57.3	64.3	68.3
Somnolence	21.4	21.4	10.5	10.5	15.8	15.8	43.8	43.8	12.6	12.6	27.6	27.6
Nausea	16.3	18.4	28.6	29.5	22.7	24.1	4.2	6.3	9.7	9.7	7.0	8.0
Thirst	14.3	15.3	1.9	1.9	7.9	8.4	20.8	21.9	1.0	1.0	10.6	11.1
Malaise	3.1	3.1	0.0	0.0	1.5	1.5	13.5	14.6	0.0	0.0	6.5	7.0
Fatigue	0.0	0.0	14.3	14.3	7.4	7.4	0.0	0.0	12.6	13.6	6.5	7.0
Constipation	14.3	15.3	0.0	0.0	6.9	7.4	12.5	12.5	1.0	1.9	6.5	7.0
Dizziness	4.1	5.1	12.4	13.3	8.4	9.4	9.4	9.4	13.6	13.6	11.6	11.6
Headache	8.2	12.2	13.3	13.3	10.8	12.8	9.4	11.5	9.7	9.7	9.5	10.6
RTI	0.0	13.3	0.0	2.9	0.0	7.9	0.0	5.2	0.0	1.0	0.0	3.0
Weight ↑	0.0	0.0	6.7	6.7	3.4	3.4	6.3	6.3	9.7	9.7	8.0	8.0
Appetite ↑	0.0	0.0	4.8	4.8	2.5	2.5	2.1	2.1	9.7	9.7	6.0	6.0
Diarrhea	6.1	8.2	8.6	8.6	7.4	8.4	1.0	2.1	3.9	3.9	2.5	3.0
SGPT ↑	2.0	2.0	0.0	0.0	1.0	1.0	8.3	8.3	0.0	0.0	4.0	4.0
Dyspepsia	7.1	7.1	1.0	1.0	3.9	3.9	1.0	2.1	0.0	0.0	0.5	1.0
Mouth dry	0.0	0.0	5.7	6.7	3.0	3.4	1.0	1.0	4.9	4.9	3.0	3.0
Hypoaesthesia	1.0	1.0	0.0	0.0	0.5	0.5	6.3	6.3	0.0	0.0	3.0	3.0
Abdominal pain	6.1	6.1	2.9	2.9	4.4	4.4	1.0	2.1	1.9	1.9	1.5	2.0
Sweating ↑	2.0	2.0	5.7	5.7	3.9	3.9	0.0	0.0	2.9	2.9	1.5	1.5
Insomnia	0.0	0.0	4.8	5.7	2.5	3.0	5.2	5.2	1.0	1.0	3.0	3.0
SGOT ↑	1.0	1.0	0.0	0.0	0.5	0.5	5.2	5.2	0.0	0.0	2.5	2.5

1) ADR = Assessment of the causal relationship by the investigator: 'Definitely', 'Probably', 'Possibly' or 'Unlikely' related.

RTI = Respiratory tract infection; ↑ = increased.

One Japanese patient treated with mirtazapine discontinued prematurely because due to increased values of transaminases, alkaline phosphatase, and creatine kinase. Plasma levels of transaminases and creatine kinase were already above normal limits at trial entry so that the relationship with trial medication remains undetermined.

All events of SGOT/SGPT increase were of mild intensity. For most of the patients with SGOT/SGPT increase, the event was reported at day 14, with all of the patients recovering during the treatment period. Such events did not lead to stopping the trial or AEs related to

hepatic disorders. For two patients (one for each treatment group), SGOT/SGPT increase was reported at the last visit of treatment period (day 42). Both patients recovered until the follow-up visit two weeks after treatment (day 56).

DISCUSSION

The clinical improvement at endpoint in patients treated with mirtazapine and fluvoxamine was similar and the primary objective of non-inferiority of mirtazapine versus fluvoxamine as assessed by $HAMD_{17}$ total score was met. The data suggest that mirtazapine has a faster onset of action than fluvoxamine. These results are also reflected by responder rates on $HAMD_{17}$ and CGI.

The comparison of efficacy on mirtazapine and fluvoxamine within the ethnic groups revealed somewhat different results for the Japanese and Caucasian population. In Japanese patients the mean antidepressant effect of mirtazapine was more pronounced, although this difference was not statistically significant. On the contrary, in Caucasian patients the mean antidepressant effect of mirtazapine appeared slightly less pronounced than that of fluvoxamine (Table 7.4). This potential interaction of efficacy with ethnicity was not statistically significant within the ITT population, but did reach statistical significance within the per-protocol population (consisting of all subjects from the ITT group without any major protocol violation, including discontinuation before two weeks of treatment).

Differences in baseline characteristics between Japanese and European patients do not offer a plausible explanation for the observed difference. First, differences are not that substantial, as can be seen in Table 7.1. Secondly, at least two characteristics that did differ have not been demonstrated predictive of efficacy, i.e., gender and weight differences.[26] Most importantly, for characteristics to explain discrepant efficacy results, these should not only be predictive of efficacy, but also be spread unevenly enough across the studied groups. Japanese patients may, for example, experience somewhat higher exposure levels due to an average difference in body weight. However, the observed difference of 12% in exposure is still not very large, and a clear dose-response relationship has never been established for either mirtazapine or fluvoxamine. If it still was to provide an explanation it should be the case that

the response in Japanese on both drugs would be higher than in Caucasians, but the reverse seems to hold.

Another, theoretical explanation for the difference in efficacy could be that mirtazapine as compared to fluvoxamine is a more potent antidepressant in Japanese than in European depressed patients. However, there is currently no known pharmacological or pharmacokinetic basis for this hypothesis. A more likely explanation seems to be that the finding is based on coincidence and that the overall results (combining both populations) present a fair estimate of the efficacy of mirtazapine as compared to fluvoxamine. It is strengthened and supported by the observation that the decrease in $HAMD_{17}$ total score on mirtazapine was similar in Japanese and Caucasian patients (-13.8 and -14.4, respectively), but improvement on fluvoxamine in Caucasian patients was higher than in Japanese patients (-15.1 and -11.7 points, respectively), and higher than documented in other studies with this compound.[4,5,27,28]

As indicated in the beginning of results parts, 12 patients were excluded from this study analysis. If we include the 12 patients, the decrease in $HAMD_{17}$ total score on mirtazapine and fluvoxamine remains similar in Japanese and Caucasian patients.

Overall, both drugs were well tolerated. Mirtazapine was associated with a higher frequency of somnolence in Japanese patients, whereas fluvoxamine was associated with a higher incidence of nausea in both Japanese and Caucasian patients.

A relatively high incidence of nausea in fluvoxamine treated patients fits the documented safety profile of this drug. In previous clinical studies, nausea being the most prominent adverse event has been reported in approximately 30-40% of patients treated with fluvoxamine.[6,29,30]

The average incidence of somnolence on mirtazapine (28%) in this study is of a similar magnitude as found in other studies with mirtazapine (11 - 35%).[21,22] For unknown reasons, somnolence is more frequently reported in Japanese patients (44%) than in Caucasian patients (13%). Also on fluvoxamine frequency of somnolence was higher in Japanese than in Caucasian patients. In a substantial number of patients experiencing somnolence, the symptom disappeared within the first two weeks.

In the study, ‘SGOT increased‘ or ‘SGPT increased’ was reported as an adverse event for only Japanese patients. The intensity of all reported ADRs (Adverse Drug Reactions: Assessment of the causal relationship by the investigator: ‘Definitely’, ‘Probably’, ‘Possibly’,

or 'Unlikely' related) were mild and no patients discontinued the drugs due to the ADRs. 'SGOT increased' or 'SGPT increased' is one of the known ADRs of mirtazapine and SSRIs, and evanescent SGOT/ SGPT increases were observed in ca 5% of Caucasian subjects in overseas studies with miratazapne.[14] No marked differences between Japanese and Caucasian patients were observed in average SGOT/SGPT levels, allowing the conclusion that changes in liver enzymes are not considered to be particular for Japanese patients.

The increased frequency of adverse events and reported transaminase increases in Japanese patients may be due to a more rigorous reporting practice in Japan, and not so much indicative of a higher vulnerability of Japanese patients to drug side effects. In particular, increases of laboratory values were typically reported as adverse events in Japan (e.g. SGOT/ SGPT increased), but were only reported in Europe when considered clinically relevant by the investigator (usually only when accompanied by symptoms like jaundice).

Although some adverse events seem to occur almost exclusively in Japanese or Caucasian subjects, the difference in safety profile between both populations might be less prominent then it seems. In some cases there seem to be adverse events which were experienced both by Japanese and Caucasian patients but which were coded differently. This is quite clear for malaise in Japanese patients and fatigue in Caucasian patients. The same phenomenon applies to thirst in Japanese patients and dry mouth in Caucasian patients. Such discrepancies may be more related to translational issues rather than reflecting a true difference in safety profile in the two ethnic groups.

Conclusions

Both mirtazapine (15-45 mg/day) and fluvoxamine (50-150 mg/day) provided an effective treatment for Japanese and Caucasian patients with major depressive disorder where mirtazapine was characterized by a faster onset of action. With respect to safety both compounds were well tolerated, revealing distinct safety profiles. Mirtazapine was characterized by increased incidences of somnolence, dry mouth/thirst and malaise/fatigue, whereas fluvoxamine is characterized by increased incidences in nausea, dyspepsia and diarrhea. In general, the safety profile of mirtazapine was comparable between Japanese and Caucasian patients. Overall there was no indication that Japanese depressed patients respond differently to mirtazapine treatment than Caucasian patients.

ACKNOWLEDGEMENTS

The Mirtazapine Bridging Study Group

Europe

P Willemse, Mouscron (Belgium); A de Nayer, Montignies-sur-Sambre (Belgium); P Schrøder, Risskov (Denmark); F Bjørndal, Greve (Denmark); O Nielsen, Holstebro (Denmark); H Wernlund, Hjørring (Denmark); J de la Gandara, Burgos (Spain); M Franco Martin, Zamora (Spain); JR Gutiérrez Casares, Badajoz (Spain); J Gailledreau, Elancourt (France); J-P Hervé, Rennes (France); M Salfati, Saint Ouen (France); O-J Høyberg, Alesund (Norway); H Rogge, Nestun (Norway); I Eikenaes, Nestun (Norway; O Kavlie, Oslo (Norway); T Dammen, Oslo (Norway); M Hompland, Bønes (Norway); G Zwartjes, Tilburg (Netherlands); D Kromdijk, Zwolle (Netherlands); R ten Kate, Oldenzaal (Netherlands); D van Hyfte, Weert (Netherlands).

Japan

Takeshi Inoue, Saporro (Hokkaido); Toshiki Shioiri, Niigata (Niigata); Tsuyoshi Kondo, Hirosaki (Aomori); Masahito Fushimi, Akita (Akita); Akihiko Takahashi, Yamato (Kanagawa); Hideo Muraoka, Sagamihara (Kanagawa); Fujio Yokoyama, Iruma-gun (Saitama); Takami Yagyu, Osaka; Yoshifumi Watanabe, Ube (Yamaguchi); Shunzo Watanabe, Hirosaki (Aomori); Yuichi Sato, Kumagaya (Saitama); Shigero Mitsuno, Hofu (Yamaguchi); Shozo Fujii, Shimonoseki (Yamaguchi); Yoshio Igarashi, Chichibu (Saitama); Ikuo Kudo, Yokohama (Kanagawa); Masashi Yasui, Sagamihara (Kanagawa); Shinano Isawa, Kawasaki (Kanagawa); Hirofumi Wakatabe, Chigasaki (Kanagawa); Yasuyuki Kawakami, Oume (Tokyo); Keiichi Kumagai, Niigata (Niigata); Kazuhiko Nakayama, Minato-ku (Tokyo); Hiroo Kasahara, Kashiwa (Chiba); Ryoichi Ikeda, Shinjyuku-ku (Tokyo); Nobuhiko Ishii, Yokohama (Kanagawa); Shun Yamazumi, Kofu (Yamanashi); Koichi Yamada, Yokohama (Kanagawa); Michiya Sugawara, Ota-ku (Tokyo); Teruhiko Higuchi, Ichikawa (Chiba); Masahiro Ishikawa, Fujisawa (Kanagawa); Masatsugu Nagao, Kure (Hiroshima); Masahiro Nankai, Chiyoda-ku (Tokyo); Noriaki Kataoka, Chitose (Hokkaido); Motofumi Fukahori, Fukuoka (Fukuoka); Kazumasa Hirabayashi, Kita Azumi-gun (Nagano).

Special thanks to Nippon Solvay for providing double-blind fluvoxamine tablets and matching placbo.

REFERENCES

1. Devane CL. Comparative safety and tolerability of selective serotonin reuptake inhibitors. *Human Psychopharmacology: Clinical and Experimental* 1995;10(S3):S185-S193.
2. Entsuah AR, Thase ME. Response and remission rates in different subpopulations with major depressive disorder administered venlafaxine, selective serotonin reuptake inhibitors, or placebo. *Journal of Clinical Psychiatry* 2001;62(11):869-877.
3. Westenberg HG. Pharmacology of antidepressants: selectivity or multiplicity? *Journal of Clinical Psychiatry* 1999;60(suppl 17):4-8.
4. Feighner JP, Boyer WF, Meredith CH, Hendrickson GG. A placebo-controlled inpatient comparison of fluvoxamine maleate and imipramine in major depression. *International Clinical Psychopharmacology* 1989;4(3):239-244.
5. Roth D, Mattes J, Sheehan KH, Sheehan DV. A double-blind comparison of fluvoxamine, desipramine and placebo in outpatients with depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 1990;14(6):929-939.
6. Wagner W, Zaborny B, Gray T. Fluvoxamine. A review of its safety profile in world-wide studies. *International Clinical Psychopharmacology* 1994;9(4):223-228.
7. De Boer T, Ruigt G, Berendsen H. The α2-selective adrenoceptor antagonist org 3770 (mirtazapine, Remeron®) enhances noradrenergic and serotonergic transmission. *Human Psychopharmacology: Clinical and Experimental* 1995;10(S2):S107-S118.
8. Haddjeri N, Blier P, de Montigny C. Effect of the alpha-2 adrenoceptor antagonist mirtazapine on the 5-hydroxytryptamine system in the rat brain. *Journal of Pharmacology and Experimental Therapeutics* 1996;277(2):861-871.
9. Dahl M-L, Voortman G, Alm C, Elwin C-E, Delbressine L, Vos R, Bogaards J, Bertilsson L. In vitro and in vivo studies on the disposition of mirtazapine in humans. *Clinical Drug Investigation* 1997;13(1):37-46.
10. Störmer E, von Moltke LL, Shader RI, Greenblatt DJ. Metabolism of the antidepressant mirtazapine in vitro: contribution of cytochromes P-450 1A2, 2D6, and 3A4. *Drug Metabolism and Disposition* 2000;28(10):1168-1175.
11. Timmer CJ, Sitsen JA, Delbressine LP. Clinical pharmacokinetics of mirtazapine. *Clinical Pharmacokinetics* 2000;38(6):461-474.
12. Bremner JD. A double-blind comparison or Org 3770, amitriptyline, and placebo in major depression. *Journal of Clinical Psychiatry* 1995.
13. Claghorn JL, Lesem MD. A double-blind placebo-controlled study of Org 3770 in depressed outpatients. *Journal of Affective Disorders* 1995;34(3):165-171.
14. Khan M. A randomised, double-blind, placebo-controlled, 5-weeks' study of org 3770 (mirtazapine) in major depression. *Human Psychopharmacology: Clinical and Experimental* 1995;10(S2):S119-S124.
15. Leinonen E, Skarstein J, Behnke K, Ågren H, Helsdingen JT, Group NAS. Efficacy and tolerability of mirtazapine versus citalopram: a double-blind, randomized study in patients with major depressive disorder. *International Clinical Psychopharmacology* 1999;14(6):329-337.
16. Mullin J, Lodge A, Bennie E, McCreadie R, Bhatt GS, Fenton G. A multicentre, double-blind, amitriptyline-controlled study of mirtazapine in patients with major depression. *Journal of Psychopharmacology* 1996;10(3):235-240.
17. Richou H, Ruimy P, Charbaut J, Delisle J, Brunner H, Patris M, Zivkov M. A multicentre, double-blind, clomipramine-controlled efficacy and safety study of org 3770. *Human Psychopharmacology: Clinical and Experimental* 1995;10(4):263-271.
18. Zivkov M, De Jongh GD. Org 3770 versus amitriptyline: A 6-week randomized double-blind multicentre trial in hospitalized depressed patients. *Human Psychopharmacology: Clinical and Experimental* 1995;10(3):173-180.
19. Van Moffaert M, De Wilde J, Vereecken A, Dierick M, Evrard J, Wilmotte J, Mendlewicz J. Mirtazapine is more effective than trazodone: a double-blind controlled study in hospitalized patients with major depression. *International Clinical Psychopharmacology* 1995.
20. Smith WT, Glaudin V, Panagides J, Gilvary E. Mirtazapine vs. amitriptyline vs. placebo in the treatment of major depressive disorder. *Psychopharmacology Bulletin* 1990;26:191-196.

21. Behnke K, Søgaard J, Martin S, Bäuml J, Ravindran AV, Ågren H, Vester-Blokland ED. Mirtazapine orally disintegrating tablet versus sertraline: a prospective onset of action study. *Journal of Clinical Psychopharmacology* 2003;23(4):358-364.
22. Benkert O, Kohnen R. Mirtazapine compared with paroxetine in major depression. *Journal of Clinical Psychiatry* 2000;61(9):656-663.
23. Wheatley DP, van Moffaert M, Timmerman L, Group M-FS. Mirtazapine: efficacy and tolerability in comparison with fluoxetine in patients with moderate to severe major depressive disorder. *Journal of Clinical Psychiatry* 1998;59(6):306-312.
24. Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. *Archives of General Psychiatry* 1988;45(8):742-747.
25. Guy W. Clinical global impression scale. *The ECDEU Assessment Manual for Psychopharmacology - Revised Volume DHEW Publ No ADM* 1976;76(338):218-222.
26. Morishita S, Arita S. Possible predictors of response to fluvoxamine for depression. *Human Psychopharmacology: Clinical and Experimental* 2003;18(3):197-200.
27. March JS, Kobak KA, Jefferson JW, Mazza J, Greist J. A double-blind, placebo-controlled trial of fluvoxamine versus imipramine in outpatients with major depression. *Journal of Clinical Psychiatry* 1990;51:200-202.
28. Norton K, Sireling L, Bhat A, Rao B, Paykel E. A double-blind comparison of fluvoxamine, imipramine and placebo in depressed patients. *Journal of Affective Disorders* 1984;7(3-4):297-308.
29. Ware MR. Fluvoxamine: a review of the controlled trials in depression. *Journal of Clinical Psychiatry* 1996;58:15-23.
30. Ottevanger E. SSRI treatment-emergent nausea. *Neuropsychopharmacology* 1994;10(Suppl 3):104S.

Chapter 8

Summary

and concluding remarks

The development of new drugs is a long-term, costly and risky process, with total development time due to a limited patent period being constantly under pressure. In the absence of pre-clinical models, which predict the effects of new substances on central nervous system disorders with reasonable certainty, the development of new drugs for disorders of the central nervous system (CNS), i.e., psychiatric and neurological diseases, is usually more difficult than for other indications. In CNS disorders, the development process is also hampered by the lack of objective criteria for determining the correct diagnosis and severity of diseases. Investigators are therefore required to interpret disease severity by observation and (standardized) interrogation of the patient, resulting in a risk of subjective assessment errors or bias. Body functions and abnormalities, such as heart rate, blood pressure, endocrine disorders, tumor size and bacterial infections, can usually be measured more objectively with relatively high accuracy or can be simply assessed through the use of biomarkers such as the concentration of certain substances in the blood. When determining the presence of depression or the seriousness of a psychosis, and when measuring treatment results in such conditions, there is still much room for interpretation. Taken together, absence of pre-clinical models, biomarkers, and objective assessment criteria slow down the development of CNS products, and may explain the relatively higher costs (on average 18%, or about 100 million euros), longer time (on average 37%, or about two years), and lower chance of success (on average 50% less, or only 1 instead of 2 out of 12 experimental products reaching the finish line) associated with the development of drugs for psychiatric versus somatic diseases. More details can be found on this in **Chapter 1**. It is therefore not surprising that the pharmaceutical industry, on the one hand, is continually seeking more efficient methods of research for the development of CNS drugs and, on the other hand, withdraws more and more from this field because the costs no longer outweigh overall revenues.

In this dissertation, some strategies for early and late stage development of antidepressants and antipsychotics are presented, all configured and tested with the aim to reduce costs and/or increase the chance of success, thereby accelerating the total development time of new products. In **Chapter 2**, the disturbing effect of placebo response on efforts to establish the efficacy of new products, and factors potentially contributing to placebo response in antidepressant trials are reviewed. If in double-blind studies, patients respond well to treatment with placebo (i.e., a substance with no specific activity), it is difficult to demonstrate that an investigational product (i.e., active ingredient) is effective against the

condition for which it is tested. A high placebo response easily leads to failure of a study, requiring it to be re-performed as long as there is belief in the product and hope for better results. In depression studies, in particular, placebo response is a significant problem, and researchers have been trying to explain and identify the main factors contributing to placebo response in this disease for many years. Usually, through meta-analysis of results from independent RCTs, calculating the overall likelihood of improvement or response of patients randomized to placebo, or trial success rate under different settings. The results of all known meta-analyses published so far are listed in Chapter 2 and compared with each other. It appears that such (inappropriately called) "meta-analyses," not rarely produce conflicting results, and provide little insight into which patient characteristics or test methods facilitate or attenuate response to placebo. Prospective studies, exploring clinical improvement in patients treated with placebo under certain conditions might be more useful for the identification of predictors of placebo response. However, such studies are extremely costly and difficult to execute because of medical and ethical dilemmas when patients are deliberately kept from effective treatment. In conclusion, after 25 years of re-analyzing previous studies, we still have to accept the fact that response to placebo is unpredictable. Luckily, for the medical treatment of patients, it does not matter what the cause of the clinical improvement is. Any contribution to a good outcome of treatment is welcome here: whether it is caused by the drug itself, by hopeful expectations, or other aspects makes no difference in clinical practice.[1]

Chapter 3 shows the practical example of an early stage, double-blind, placebo-controlled study in which, under most stringent test conditions and with sophisticated analysis techniques, as little as possible is left to chance to show that treatment augmentation with a new drug is effective against the (often difficult to measure) negative symptoms of schizophrenia (such as retardation, apathy, and the inability to express or recognize emotions). The usage of (*1*) second opinion, (*2*) separate assessment scales for allowing suitable patients to the study on the one hand, and measuring clinical efficacy on the other, as well as (*3*) analysis methods whereby primary treatment effects are discriminated from secondary effects, helps to provide a clear answer to the question as to whether treatment augmentation is useful or not. The glycine uptake inhibitor with the codename Org 25935 barely distinguished itself from placebo in this trial, so that further development of this product was deemed pointless and the project abandoned. A competitive company, which examined a compound with identical action mechanisms, obtained similar results (i.e., a mild, almost statistically

significant improvement on placebo at the lowest dose tested) in parallel.[2] However, their studies were devoid of aforementioned, sophisticated techniques, so that stakeholders in that company decided it was worth to embark on the subsequent conduct of large-scale Phase III studies. After investment of an estimated $100 million, there was still no convincing evidence of the product's effectiveness collected, so that further research was discontinued.[3,4] This shows that the implementation of advanced methods, although costly, can be of great value for a timely discontinuation of fruitless development programs, especially when investigational products exhibit a novel mechanism of action.

In **Chapter 4**, results from a large-scale, Phase III study are described, conducted under more or less 'naturalistic' conditions, whereby most patients are allowed to participate when meeting appropriate diagnostic criteria. In earlier research phases, patients are frequently excluded from study participation, especially when using certain (legal or illegal) psychotropic drugs, suffering from a chronic disease or infection, or any other condition that might put a participant at increased risk for drug interactions or side effects, compromise compliance, or complicate the collection of clean data and interpretation of research results. The advantage of a naturalistic study is that its results can be regarded as more representative for daily practice than when large groups of patients had to be excluded. On the other hand, it may also be detrimental to the sponsor of the study if the results appear to be less successful than expected. We see this happening in the study in Chapter 4, where the efficacy and tolerability of asenapine seem somewhat less promising in comparison with those of the market leading product olanzapine (except for the relatively high risk of weight gain on the latter). It is not surprising that reviewers, key opinion leaders and therapists look at the results of such large studies very critically, and are inclined to adhere more value to these than to the so-called pivotal, placebo-controlled studies, which are usually performed on a significantly smaller scale in more or less 'clean' patients.[5] Finally, a naturalistic study may be an additional valuable source of information for the identification of potential factors predisposing to poor response or increased vulnerability to side effects, such as severity and duration of disease prior to treatment, young or high age, drug abuse, etc..

In order to meet the demand for continued use of a promising investigational product after a patient's participation in a study, it is not unusual to carry out a so-called humanitarian treatment continuation according to the protocol after completion of the main study period. An additional advantage of such an approach is that patients are more likely to participate in

the main study, facilitating the conduct of extensive studies within a restricted time frame. The alternative, allowing treating physicians compassionate use of a new remedy prior to registration for patients who respond poorly to already available medication, is less attractive due to regulatory, medical-ethical, and market-strategic constraints. After all, with compassionate use potential side effects will be collected under poorly controlled conditions, and physicians may get biased against the product before the new drug is placed on the market. **Chapter 5** summarizes trends observed in a humanitarian extension of a study that was earlier presented in Chapter 4. It is interesting to see that the percentage of patients withdrawing from the study in both treatment groups is comparable, although in the asenapine group more adverse reactions are reported, and more comedication is prescribed. Also interesting to see is that the on average 5 kg increase in body weight within the olanzapine group appears irreversible, whereas body weight increase in the asenapine group is not only much more limited (on average 1 kg) but even shows a trend to disappear during prolonged treatment. Standardized documentation of this type of data would not have been possible on the basis of compassionate use and is relevant to any physician detecting first signs of weight gain in a patient after treatment start. After all, with long-term use of antipsychotics, metabolic disorders that result in weight gain are an important risk factor for diabetes and secondary cardiovascular problems.[6]

Study drop-out as a result of poor efficacy, tolerability, or compliance, should always be anticipated and compensated for by adequate sample sizes. Drop-out tends to be high in clinical trials involving patients with schizophrenia, who may withdraw their consent for voluntary participation at any moment. As a result, not only measures to increase the motivation of patients to participate but also precautions to avoid drop-out are important means to enhance the efficiency of studies and reduction of development costs. The latter requires insight into the primary drivers of patient satisfaction with experimental treatment. In **Chapter 6**, a post-hoc analysis of patient data from the in Chapter 4 reported study seeks to provide an answer to the question why many participants in schizophrenia studies withdraw prematurely, often despite initial signs of good response. Large-scale, naturalistic studies are particularly useful for this kind of exploratory research because, unlike in a sample of independent studies, all patients receive the same treatment under similar conditions. Unfortunately, no explanation could be found for this phenomenon, so that study continuation by schizophrenia patients remains unpredictable and difficult to encourage.

When a sponsor seeks market expansion for a currently registered product within the region where the guidelines of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) are endorsed, one can rely on the E5 Directive on Acceptance Of foreign data.[7] Ideally, records of results obtained elsewhere can be used to apply for registration in the new region, provided it is made likely (on basis of a bridging study) that ethnic differences will not influence the efficacy and safety profile established elsewhere. This prevents unnecessarily high investment and repetition of steps. A bridging study was conducted with the antidepressant Remeron® (mirtazapine) to obtain registration in Japan after the drug had been registered in the United States and Europe. The study was primarily designed to demonstrate non-inferiority of mirtazapine at doses varying between 15 and 45 mg/day in comparison with fluvoxamine at doses between 50 and 150 mg/day in Japanese and Caucasian patients with major depressive disorder. The results are described in **Chapter 7**. A bridging study is successful when the effectiveness and adverse event profile of the test product at particular dose level(s) overlap with the drug's efficacy & safety profile that was earlier established outside the new region. Although clinical improvement on mirtazapine was at least as good in Japanese patients as in Caucasians, albeit at the expense of a somewhat higher incidence of somnolence, Japanese authorities were not immediately convinced that an identical efficacy & safety profile of Remeron® would apply across the ethnic regions, and requested an additional placebo-controlled study with mirtazapine at fixed dose levels. A placebo-controlled, fixed-dose trial had been among the initial suggestions for a bridging study, but Japanese co-workers feared at that time it would be difficult to recruit patients and/or ethic committees would never approve a study in which part of the patients would be kept from adequate treatment. This was an expensive judgment error, causing two years of delay in trial registration. Although not stipulated as absolute requirements, it shows that in bridging studies it is always preferable to test at fixed doses and randomize at least a minority of the patients to placebo to serve as control group.

The studies included in this thesis are just a small grip from the arsenal on possibilities for drug research in psychiatry to be as efficient as possible. The central goal in most clinical trials is to demonstrate efficacy and response to a trial drug. However, it is often non-response that is easier to determine, as it is not always clear when a psychiatric patient feels "better". Although drug treatment is primarily aimed at suppressing disease symptoms, patients often

have hopeful expectations of an improved quality of life at the onset of pharmacological treatment. In order to ensure that doctors and patients have the same view of the need for and success of treatment, it is important that clinical investigators clearly explain the potential benefits and risks of study participation before asking a patient's informed consent. At the same time, patients should be made aware of the fact that a certain level of responsibility and sincerity is expected from them as a 'partner in science'. Till now, investigators are hardly trained in adequate or most appropriate methods for obtaining informed consent from their patients. More than any other strategy, this could contribute to the participation of motivated but also critical patients in clinical trials, the collection of representative and more complete datasets, and the increase of drug research efficiency in psychiatric diseases.

Successful and timely completion of clinical studies depends on many factors, including the availability of investigators and patients, their enthusiastic willingness to participate, and intensive medical monitoring and follow-up during the conduct of a trial. Since expectations among investigators and patients are difficult to control for in an experimental setting, double-blind randomization is essential. Blinding forms the basis for a fair, and adequate comparison of treatment regimes. However, enthusiasm about the promising profile of a new drug, and hope for improvement during its use should never be eliminated where it belongs: in clinical practice. Given the relatively high response rate on placebo in psychiatric illnesses and vain search for its proximate causes, as well as the more robust criteria for 'non-response' or 'remission' versus 'response', the field might benefit from future efforts directed towards the identification of factors predicting *poor* response on placebo, or (fast) *remission* on active treatment.

Epilogue

As with most trials conducted within the context of drug development programs, the studies presented in **Chapters 3, 4, 5,** and **7** have no other aim than to provide testimony of the experimental drugs' potential to be further developed and brought to market. Since these studies involve three different drugs that are tested in three different indications, it is difficult to tie the results and evaluate how these may contribute to scientific or clinical progress. The same for the outcome variables 'placebo response' and 'drop-out' (**Chapters 2** and **6**) that were explored in two different ways, and in two different indications (depression and schizophrenia). What latter two variables do have in common is that both may have

implications for trial failure. However, main reasons for trial failure may differ considerably between depression and schizophrenia (although this topic is not well explored, to the best of my knowledge). The fact that no clear contributors to placebo response (in depression) or patient withdrawal (in schizophrenia) could be identified is not really helpful either when we want to summarize the overall merits of our work.

In some way or another, what the studies presented in this thesis do have in common, is that the results of none of those appear to have been capable in changing daily clinical practice or reforming common approaches taken in drug development. As said, I was unable to identify predictors of placebo response in depression or patient withdrawal in schizophrenia (**Chapters 2** and **6**). If one or more predictors would have been found, these would certainly be taken into consideration when new studies are to be designed in these indications. Despite a meticulous study design and conduct (**Chapter 3**), including an innovative second-opinion evaluation of clinical symptoms and advanced statistics, I was unable to identify a significant effect of Org 25935 on negative symptoms in schizophrenia; the small effect that was observed appeared to be mainly non-specific, i.e., secondary to changes in positive, general psychopathological, depressive, neurocognitive, and extrapyramidal symptoms, and was also therefore unimportant. If a specific, statistically significant effect on negative symptoms would have been found, it is not unlikely that the innovative, expert-rater assisted score evaluation technique that was for the first time used in this trial, would have become more broadly adopted in future psychiatric drug trials to reduce bias. Especially, when post-hoc evaluations would have shown that effect sizes increased after raters adjusted their preliminary scores on the basis of second opinion. In addition, path analytic approaches could be more frequently added to the toolbox of drug developers and included in the primary analysis of trial results, when it had been demonstrated that the effect on outcome was more or less specific in the symptom domain of interest. Unfortunately, in the absence of any significant effect, the long-term merits of the innovative study approach remain speculative.

The results from the studies presented in **Chapters 4, 5** and **7** demonstrated that experimental treatment with asenapine and mirtazapine was non-inferior to treatment with the established drugs olanzapine and fluvoxamine, in schizophrenia and major depression respectively. On the other hand, the studies were at the same time not capable of demonstrating any meaningful improvement over the established products and did not provide new insights into optimized treatment in these psychiatric diseases.

Nevertheless, despite a genuine lack of potentially influential results, these studies deserve as much attention as those with a more positive outcome. First of all, although the potential relevance of the glutamatergic axis in the expression of schizophrenia symptoms appeared firmly established, the fact that two companies in row were unsuccessful in demonstrating the validity of the concept of GlyT-1 inhibition, is an important warning signal that is worth to note.[8,9 10-13] The sharing of such negative results through publication may keep a third from exposing still more patients in vain to experiments with similar, ineffective drugs. Secondly, based on short- and long-term treatment results presented in **Chapters 4, 5,** and **7**, both asenapine and mirtazapine could be regarded as useful alternatives for patients experiencing unacceptable side-effects from other prescribed drugs.

The fact that I was not able to identify strong predictors of placebo response or patient withdrawal in depression and schizophrenia, makes it not advisable (at this stage) to try tweaking the most commonly applied enrolment criteria for clinical trials towards enhanced signal detection or trial completion rates, solely on the basis of personal judgment or preliminary results from published meta-analyses and surveys in these fields. The current lack of evidence, however, does not necessarily imply that the conduct of meta-analyses for detection of new leads in trial optimization is inappropriate and to be entirely abandoned. In a recent paper, Cipriani and co-workers (2018) note that future research could seek to combine aggregate and individual-patient data from RCT in so-called network meta-analyses.[14] They express the belief that such analyses might be better suited than currently applied methods to allow the prediction of personalized clinical outcomes, including the estimate of comparative efficacy or risk for premature discontinuation. There is a common notion that the effect sizes of antidepressants and antipsychotics in RCTs are rather modest, suggesting that these drugs as a whole are barely more effective than placebo.[14,15] Previous work has suggested that treatment-naïve patients in their first episode of schizophrenia respond better than chronically ill (i.e., earlier treated) patients.[16] A similar trend was observed in the naturalistic study that we conducted (Chapter 4, data not shown). This holds promises for RCTs involving patients in their first episode only, which might show higher effect sizes of drugs versus placebo (at least in schizophrenia) than RCTs involving more chronically ill patients. On the other hand, first-episode patients may be hard to find or motivate for trial participation, so that this population could be of particular interest only when new drugs are tested at relatively small-scale. That is usually the case in early development. However, it is exactly then, when

evidence of efficacy has not yet been proven, and first-episode patients are hardest to enroll in a clinical trial. Thus, although improved patient selection could, in theory, be helpful in drug development, the field might still in the future benefit more from integrated data processing methods like the network meta-analysis mentioned above. Open access to the data of sponsored RCTs may be the most critical prerequisite of the latter approach to become successful. The identification of reliable response predictors will allow trials to become designed as efficiently as possible to demonstrate the effectiveness of new treatments. Although certainly contributing to the replicability of studies and cost savings, whether better insight in response mediators can also meaningfully shorten the drug development trajectory in neuropsychiatric diseases remains doubtful. That goal may be just too pretentious, as long as there are no objective criteria and/or biomarkers for the accurate diagnosis and assessment of disease severity.

REFERENCES

1. Schoemaker JH. Benefits of placebos. *Nature* 1995;377(6545):98-98.
2. Umbricht D, Alberati D, Martin-Facklam M, Borroni E, Youssef EA, Ostland M, Wallace TL, Knoflach F, Dorflinger E, Wettstein JG, Bausch A, Garibaldi G, Santarelli L. Effect of bitopertin, a glycine reuptake inhibitor, on negative symptoms of schizophrenia: a randomized, double-blind, proof-of-concept study. *JAMA Psychiatry* 2014;71(6):637-646.
3. Bugarski-Kirola D, Iwata N, Sameljak S, Reid C, Blaettler T, Millar L, Marques TR, Garibaldi G, Kapur S. Efficacy and safety of adjunctive bitopertin versus placebo in patients with suboptimally controlled symptoms of schizophrenia treated with antipsychotics: results from three phase 3, randomised, double-blind, parallel-group, placebo-controlled, multicentre studies in the SearchLyte clinical trial programme. *The Lancet Psychiatry* 2016;3(12):1115-1128.
4. Bugarski-Kirola D, Wang A, Abi-Saab D, Blättler T. A phase II/III trial of bitopertin monotherapy compared with placebo in patients with an acute exacerbation of schizophrenia–results from the CandleLyte study. *European Neuropsychopharmacology* 2014;24(7):1024-1036.
5. Siu J. New magic or old tricks? A psychiatry update. *CSHP BC Branch Clinical Symposium, Vancouver (BC)* 2013.
6. Newcomer JW. Antipsychotic medications: metabolic and cardiovascular risk. *Journal of Clinical Psychiatry* 2007;68:8-13.
7. ICH. International Conference on Harmonization Tripartite Guidance E5: Ethnic Factors in the Acceptability of Foreign Data. *The US Federal Register* 1998;63(111):31790–31796.
8. Hons J, Zirko R, Ulrychova M, Cermakova E, Doubek P, Libiger J. Glycine serum level in schizophrenia: relation to negative symptoms. *Psychiatry Research* 2010;176(2):103-108.
9. Sur C, Kinney GG. The therapeutic potential of glycine transporter-1 inhibitors. *Expert Opinion on Investigational Drugs* 2004;13(5):515-521.
10. Gomeza J, Ohno K, Betz H. Glycine transporter isoforms in the mammalian central nervous system: structures, functions and therapeutic promises. *Current Opinion in Drug Discovery & Development* 2003;6(5):675-682.
11. Hutson P, Clark J, Cross A. CNS target identification and validation: avoiding the valley of death or naive optimism? *Annual Review of Pharmacology and Toxicology* 2017;57:171-187.
12. Kantrowitz JT. Managing Negative Symptoms of Schizophrenia: How Far Have We Come? *CNS Drugs* 2017;31(5):373-388.
13. Cioffi CL. Glycine transporter-1 inhibitors: a patent review (2011-2016). *Expert Opinion on Therapeutic Patents* 2018(just-accepted).
14. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JP. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet* 2018.
15. Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B, Samara M, Rabaioli M, Bächer S, Cipriani A. Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, bayesian meta-analysis, and meta-regression of efficacy predictors. *American Journal of Psychiatry* 2017;174(10):927-942.
16. Lieberman J, Jody D, Geisler S, Alvir J, Loebel A, Szymanski S, Woerner M, Borenstein M. Response in First-Episode Schizophrenia. *Archives of General Psychiatry* 1993;50:369-376.

APPENDIX A
Nederlandse samenvatting

De ontwikkeling van nieuwe geneesmiddelen is een langdurig, kostbaar en riskant proces, waarbij de totale ontwikkelingstijd als gevolg van een beperkte patentduur constant onder druk staat. Door het ontbreken van pre-klinische modellen, aan de hand waarvan de effecten van nieuwe stoffen op stoornissen in het centrale zenuwstelsel met redelijke zekerheid kan worden voorspeld, is de ontwikkeling van nieuwe geneesmiddelen voor psychiatrische en neurologische ziekten doorgaans nog lastiger dan die voor andere indicaties. Het ontwikkelingsproces wordt bij deze stoornissen met name ook bemoeilijkt door het ontbreken van objectieve criteria voor het vaststellen van de juiste diagnose en ernst van een aandoening. Onderzoekers zijn derhalve genoodzaakt om aan de hand van observatie en (al dan niet gestandaardiseerde) ondervraging van de patiënt zich een beeld te vormen van het ziektebeeld, met als gevolg een risico op subjectieve inschattingsfouten. Dit in tegenstelling tot het objectief meten van allerlei lichaamsfuncties en afwijkingen, zoals hartslag, bloeddruk, endocriene stoornissen, tumorgrootte en bacteriële infecties, die redelijk eenvoudig en met grote nauwkeurigheid kunnen worden gemeten of afgeleid van de concentratie van bepaalde stoffen in het bloed. Bij het vaststellen van een depressie of ernst van een psychose, en bij het meten van behandelingsresultaten van dergelijke aandoeningen, resteert nog altijd veel ruimte voor interpretatie. Bij elkaar genomen zijn deze ongewisse factoren er mede de oorzaak van dat bij de ontwikkeling van geneesmiddelen tegen ziekten van het centrale zenuwstelsel de klinische fase gemiddeld 18% (ca. 100 miljoen euro) duurder is, 37% (ca. twee jaar) langer duurt, en 50% minder kans van slagen heeft (slechts 1 op de 12 stoffen haalt de eindstreep) dan bij de ontwikkeling van geneesmiddelen tegen somatische ziekten. Details hierover zijn terug te vinden in **Hoofdstuk 1**. Het is dan ook niet verwonderlijk dat de farmaceutische industrie voor de ontwikkeling van geneesmiddelen, die inwerken op het centrale zenuwstelsel, enerzijds voortdurend op zoek is naar meer efficiënte methoden van onderzoek

en, anderzijds, zich meer en meer terugtrekt uit dit gebied omdat de kosten niet langer opwegen tegen de baten.

In dit proefschrift komen successievelijk een aantal strategieën aan bod die zijn bedacht en getest gedurende de vroege en late ontwikkelingfasen van antidepressiva en antipsychotica, alle met het oogmerk om de kosten te drukken, de kans op succes te vergroten, en daarmee het ontwikkelingstraject te bespoedigen. In **Hoofdstuk 2** wordt vooraleerst stilgestaan bij het op grote schaal optreden van placebo-respons en de achterliggende redenen hiervoor. Als in dubbelblind onderzoek veel patiënten goed reageren op behandeling met een placebo (stof zonder specifieke werking) is het lastig om aan te tonen dat een proefmiddel (met actief bestanddeel) werkzaam is tegen de aandoening waarvoor het wordt getest. Een hoge placebo-respons leidt dan al gauw tot het falen van een studie, die dan opnieuw moet worden uitgevoerd met de hoop op beter resultaat. Met name in depressie-studies vormt placebo-respons een omvangrijk probleem, en er wordt al jaren door middel van zogenaamde meta-analyses gezocht naar factoren die placebo-respons beïnvloeden. De resultaten van alle bekende meta-analyses zijn in Hoofdstuk 2 op een rij gezet en met elkaar vergeleken. Daaruit blijkt dat de 'meta-analyses', zoals die zijn uitgevoerd, in strikte zin niet eens meta-analyse mogen worden genoemd, in vele gevallen tegenstrijdige resultaten opleveren, en weinig inzicht verschaffen in de patiëntkenmerken of proefmethoden, die de respons op placebo faciliteren of juist beteugelen. Prospectieve studies, die het optreden van een placebo-respons gericht onderzoeken in subpopulaties van patiënten die met placebo zijn behandeld onder bepaalde condities, zouden hier mogelijk een antwoord op kunnen geven. Echter, dergelijke studies zijn uiterst kostbaar en lastig uitvoerbaar, omdat er medisch-ethische dilemma's kleven aan het patiënten bewust onthouden van effectieve medicatie. Concluderend kan worden gesteld dat men zich, na 25 jaar her-analyseren van oude studies, voorlopig nog zal moeten neerleggen bij het gegeven dat respons op placebo onvoorspelbaar is, in plaats van het ongebreideld blijven navorsen van factoren die hier mogelijk toe kunnen bijdragen. Voor de medische behandeling van patiënten maakt het sowieso niet uit wat de oorzaak van de klinische verbetering is. Elke bijdrage aan een goed resultaat van de behandeling is hier welkom: of die nu door het middel zelf wordt veroorzaakt, door hoopvolle verwachtingen, of door andere aspecten, maakt geen verschil in de klinische praktijk.[1]

Hoofdstuk 3 geeft het praktijkvoorbeeld van een vroege fase, dubbelblinde, placebo-gecontroleerde studie waarin, na te zijn opgetuigd met verscherpte proefmethoden en

analysetechnieken, zo weinig mogelijk aan het toeval wordt overgelaten om aan te tonen dat toevoeging van een nieuw middel op de bestaande behandeling werkt tegen de (veelal lastig herkenbare en te meten) negatieve symptomen van schizofrenie (zoals teruggetrokkenheid, apathie, en het gebrekkige vermogen om eigen emoties uit te drukken of die van anderen te herkennen). Door het gebruik van (*1*) second opinion, (*2*) aparte beoordelingschalen voor het toelaten van geschikte patiënten tot de studie enerzijds, en het meten van effecten anderzijds, en (*3*) opsplitsing van primaire en secundaire effecten van behandeling tijdens de analyse, wordt een eenduidig antwoord verkregen op de vraag of aanvulling op de bestaande behandeling met het nieuwe testmiddel zin heeft of niet. De resultaten laten zien dat de glycine opname-remmer met codenaam Org 25935 zich niet of nauwelijks onderscheidt van placebo, zodat verdere ontwikkeling van dit middel zinloos werd geacht. Een concurrerend bedrijf, dat een stof met eenzelfde werkingsmechanisme onderzocht, verkreeg in diezelfde periode vergelijkbare resultaten (zijnde een lichte, bijna statistisch significante verbetering op placebo met de laagste, geteste dosis).[2] Echter, aangezien het in hun studie ontbrak aan verfijnde technieken en mogelijkheden ter vergelijking, zoals gebruikt in de voorbeeldstudie, werd door dit bedrijf wel besloten tot grootschalig fase III onderzoek. Na investering van een naar schatting 100 miljoen dollar bleek de stof alsnog niet te werken en werd verder onderzoek gestaakt.[3,4] Dit maakt aannemelijk dat de introductie van geavanceerde meetmethoden in de vroege ontwikkeling van een stof weliswaar prijzig is maar van grote waarde kan zijn om, later, dure beslissingsfouten te voorkomen, vooral wanneer er sprake is van een totaal nieuw werkingsmechanisme van het te testen middel.

In **Hoofdstuk 4** wordt verslag gedaan van een grootschalige fase III studie onder min of meer 'naturalistische' condities, waarbij patiënten die voldoen aan de diagnostische criteria vrijwel allemaal zijn toegestaan om deel te nemen. In vroege onderzoeksfasen worden patiënten nogal eens van studiedeelname uitgesloten indien zij, bijvoorbeeld, bepaalde (al dan niet verdovende) middelen gebruiken of een chronische ziekte of infectie onder de leden hebben, die het verzamelen van zuivere gegevens en de interpretatie van onderzoeksresultaten kunnen bemoeilijken. Het voordeel van een naturalistische studie is dat de resultaten eerder representatief mogen worden geacht voor de dagelijkse praktijk dan wanneer grote groepen patiënten van deelname zijn uitgesloten. Anderzijds kan het ook nadelig zijn voor de sponsor van de studie indien de resultaten minder succesvol lijken te zijn dan gehoopt. We zien dit gebeuren bij de studie in Hoofdstuk 4, waar de effectiviteit en verdraagbaarheid van

asenapine enigszins achterblijft bij die van olanzapine (hoewel veel patiënten fors in gewicht toenemen op laatstgenoemd middel), op dat moment een eerste keusmiddel bij de behandeling van schizofrenie. Het is dan ook niet verwonderlijk dat recensenten, richtlijnbepalers en behandelaars de resultaten van dit soort studies zeer kritisch bekijken en er soms meer waarde aan hechten dan aan die van zogenaamde 'pivotal', placebo-gecontroleerde studies, welke doorgaans op beduidend kleinere schaal zijn uitgevoerd in min of meer 'schone' patiënten.[5] Tot slot kan een naturalistische studie een belangrijke bron van informatie zijn voor het natrekken van verhoogde gevoeligheid voor bijwerkingen of kans op voorspoedige genezing bij patiënten met bepaalde basiskenmerken zoals ernst en duur van ziekte voor behandeling, jong of hoge leeftijd, drugsmisbruik, etc..

Om tegemoet te komen aan de eventuele wens van patiënten om blijvend gebruik te kunnen maken van een veelbelovend product na deelname aan een studie, is het niet ongebruikelijk om een zogenaamde humanitaire voortzetting van behandeling volgens protocol uit te voeren na afsluiting van de hoofdstudie. Bijkomend voordeel van een dergelijke aanpak is dat patiënten eerder geneigd zijn om aan de hoofdstudie deel te nemen, zodat ook grote studies praktisch uitvoerbaar worden binnen een acceptabel tijdsbestek. Het alternatief, om behandelaars door middel van "compassionate use" (dat wil zeggen, op artsenverklaring) in de gelegenheid te stellen nog vóór registratie een nieuw middel te laten gebruiken bij patiënten die slecht reageren op reeds beschikbare medicatie, is grotendeels in onbruik geraakt vanwege medisch-ethische en marktstrategische bezwaren. Immers, eventuele bijwerkingen worden dan onder slecht controleerbare omstandigheden verzameld en behandelaars kunnen al voordat het nieuwe middel op de markt komt teleurgesteld zijn in bij deze specifieke groep behaalde resultaten. De resultaten van een humanitaire extensie op de studie in Hoofdstuk 4 zijn beschreven in **Hoofdstuk 5**. Het is interessant om te zien dat het percentage patiënten dat zich terugtrekt uit de studie in beide behandelingsgroepen vergelijkbaar is, ofschoon binnen de asenapine-groep meer bijwerkingen worden gerapporteerd en meer comedicatie wordt voorgeschreven. Ook interessant om te zien is dat de gemiddelde gewichtstoename binnen de olanzapine-groep bij voortgezet gebruik gedurende meer dan twee jaar op hetzelfde, hoge niveau blijft (ca. 5 kg) terwijl die binnen de asenapine-groep beperkt blijft (ca. 1 kg) en zelfs na verloop van tijd lijkt te verdwijnen. Gestandaardiseerde documentatie van dit soort gegevens zou niet mogelijk zijn geweest op basis van 'compassionate use', en is relevant voor iedere behandelaar die de eerste tekenen

van gewichtstoename bij een patiënt bespeurt na start van medicatie. Immers, bij langetermijngebruik van antipsychotica zijn stofwisselingsstoornissen die uitmonden in gewichtstoename een belangrijke risicofactor voor diabetes en secundaire cardiovasculaire problemen.[6]

Voortijdige stopzetting van studiedeelname als gevolg van slechte werkzaamheid, verdraagbaarheid of therapietrouw, moet altijd worden geanticipeerd en gecompenseerd door een adequate steekproefgrootte. Patiëntuitval is vaak hoog in klinische studies met schizofreniepatiënten, die hun toestemming voor vrijwillige deelname op elk moment kunnen intrekken. Dientengevolge zijn niet alleen maatregelen die patiëntmotivatie voor studiedeelname versterken, maar ook voorzorgsmaatregelen die patiëntuitval kunnen voorkomen belangrijke middelen om de efficiëntie van studies te verhogen en algehele ontwikkelingskosten te drukken. Dergelijke voorzorgsmaatregelen kunnen worden getroffen wanneer men inzicht heeft in de belangrijkste drijfveren van patiënttevredenheid met experimentele behandeling. In **Hoofdstuk 6** werd met een post hoc analyse van patiëntgegevens uit de in Hoofdstuk 4 gerapporteerde studie beoogd een antwoord te verkrijgen op de vraag waarom veel patiënten met schizofrenie zich zonder opgaaf van redenen voortijdig terugtrekken uit een studie waaraan zij aanvankelijk wensten deel te nemen, vaak ondanks eerste tekenen van een goede respons. Grootschalige, naturalistische studies lenen zich bij uitstek voor dit soort exploratief onderzoek omdat, anders dan bij een verzameling van op zichzelf staande studies, alle patiënten exact dezelfde behandeling krijgen onder dezelfde condities. Helaas kon hiervoor geen verklaring worden gevonden en blijft de welwillendheid van schizofreniepatiënten om deel te nemen aan onderzoek en gecontroleerde behandeling onvoorspelbaar en moeilijk te bestendigen.

Wanneer een sponsor marktuitbreiding zoekt van een inmiddels geregistreerd product binnen de regio waar de richtlijnen van de International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) worden onderschreven, kan men zich beroepen op de E5 richtlijn met betrekking tot acceptatie van buitenlandse data.[7] In het gunstigste geval kan het volledige dossier van elders verkregen resultaten worden gebruikt om registratie aan te vragen in de nieuwe regio, mits het op basis van een overbruggingsstudie aannemelijk is dat etnische verschillen geen invloed hebben op de te verwachten effectiviteit en verdraagbaarheid van het testmiddel. Dit voorkomt een enorme investering en herhaling van zetten. Een dergelijke overbruggingsstudie is uitgevoerd

met het antidepressivum Remeron® om registratie te verkrijgen in Japan nadat het middel reeds was geregistreerd in Amerika en Europa; de resultaten hiervan worden beschreven in **Hoofdstuk 7**. De voornaamste vereiste is dat een overbruggingsstudie laat zien hoe in de nieuwe regio met een bepaalde dosering(en) een vergelijkbare effectiviteit en verdraagbaarheid wordt bereikt als elders. Ofschoon de effectiviteit en verdraagbaarheid van Remeron® in Japanners niet substantieel afweek van die in Caucasiërs waren de Japanse autoriteiten hier niet onmiddellijk van overtuigd en verzochten zij om een additionele placebo-gecontroleerde studie met vaste doseringen. Dit was aanvankelijk ook de bedoeling geweest maar er bestond in eerste instantie grote weerstand onder Japanse collega's van Organon om een dergelijk onderzoek uit te voeren. Een dure inschattingsfout die laat zien dat het bij het ontwerpen van een overbruggingsstudie de voorkeur verdient om met vaste doseringen te testen en op zijn minst een bescheiden placebo-groep als controle mee te nemen.

De in dit proefschrift opgenomen studies zijn slechts een kleine greep uit het arsenaal aan mogelijkheden om geneesmiddelonderzoek in de psychiatrie zo efficiënt mogelijk te laten verlopen. Het centrale doel bij de meeste klinische studies is het aantonen van effectiviteit en respons op een proefmedicijn. Echter, het is vaak non-respons dat zich gemakkelijker laat definiëren aangezien het bij psychiatrische ziekten niet altijd duidelijk is wanneer een patiënt zich 'beter' voelt. Hoewel het gebruik van medicijnen er primair op is gericht om ziektenymptomen te onderdrukken, koesteren patiënten vaak tevens hoopvolle verwachtingen op een verbeterde kwaliteit van leven bij aanvang van een behandeling. Teneinde te bereiken dat artsen en patiënten een zelfde opvatting hebben over de noodzaak tot, en succes van een behandeling is het belangrijk dat een onderzoeker duidelijk de mogelijke voordelen en risico's van studiedeelname aan een patiënt uitlegt voordat deze daartoe vrijwillig besluit. Maar daarbij moet tevens benadrukt worden dat van een patiënt als 'partner in de wetenschap' een zekere mate van verantwoordelijkheid en oprechtheid wordt verwacht. Een aanpak die tot nu toe grotendeels onbenut is gebleven binnen het klinisch onderzoek, is het onderrichten van onderzoekers met betrekking tot een daarop toegespitste, gestandaardiseerde procedure voor het verkrijgen van een patiënten-bewustzijnsverklaring bij aanvang van een studie. Meer dan enig andere methode zou dit kunnen bijdragen tot studiedeelname van gemotiveerde maar tegelijk ook kritische patiënten, het verzamelen van representatieve en meer volledige

datasets, en het verhogen van de efficiëntie van geneesmiddelenonderzoek bij psychiatrische patiënten.

Een succesvolle en tijdige afronding van de in dit proefschrift opgenomen, klinische studies hangt af van vele factoren, waaronder de beschikbaarheid van onderzoekers en patiënten, hun enthousiaste bereidheidwilligheid om deel te nemen, intensieve medische begeleiding tijdens de uitvoer van het onderzoek, en een goede follow-up. Aangezien verwachtingen bij onderzoekers en patiënten moeilijk zijn te controleren in een experimentele setting is het principe van dubbelblinde randomisatie een uitkomst. Blindering vormt de basis voor een eerlijke, adequate vergelijking van behandelingsregiems, maar het enthousiasme over een nieuw geneesmiddel en de hoop op verbetering behoort nimmer te worden weggenomen daar waar het thuishoort: in de klinische praktijk. Gezien de relatief hoge respons op placebo in psychiatrische studies en het vergeefs zoeken naar factoren die hiertoe bijdragen, in combinatie met het bestaan van harde criteria voor de vaststelling van non-response of genezing versus verbetering, kan het lonen om in de toekomst met meer nadruk te zoeken naar voorspellende factoren van een *slechte* respons op placebo en (voorspoedige) *genezing* bij actieve behandeling.

Epiloog

Zoals met de meeste onderzoeken die worden uitgevoerd in het kader van programma's voor de ontwikkeling van geneesmiddelen, hebben de studies die worden gepresenteerd in **Hoofdstuk 3, 4, 5** en **7** geen ander doel dan om te getuigen van het potentieel van de testmiddelen om verder te worden ontwikkeld en op de markt te worden gebracht. Aangezien met deze studies drie verschillende stoffen worden getest in drie verschillende indicaties, is het lastig om de resultaten te koppelen en te evalueren in hoe verre deze kunnen bijdragen aan de wetenschappelijke of klinische vooruitgang. Hetzelfde geldt voor de uitkomstvariabelen 'placebo-respons' en 'drop-out' (**Hoofdstuk 2** en **6**) die op niet-identieke wijze werden onderzocht bij twee verschillende indicaties (depressie en schizofrenie). Wat de laatste twee variabelen in theorie met elkaar gemeen hebben, is dat beide gevolgen kunnen hebben voor het falen van gecontroleerde studies. De belangrijkste redenen voor het mislukken van gecontroleerde studies kunnen echter aanzienlijk verschillen in depressie en schizofrenie (hoewel dit onderwerp niet uitgebreid is onderzocht, voor zover ik weet). Het feit dat er geen factoren konden worden aangewezen die duidelijk invloed hebben op de mate van placebo-

respons of terugtrekking van patiënten, is ook niet echt nuttig als we de algehele verdiensten van ons werk willen samenvatten.

Op de een of andere manier hebben de studies die in dit proefschrift worden gepresenteerd met elkaar gemeen, dat de resultaten van geen van deze zullen leiden tot hervorming van de dagelijkse klinische praktijk of geijkte onderzoeksmethoden bij de ontwikkeling van nieuwe geneesmiddelen. Zoals gezegd, was ik niet in staat om een vinger te krijgen achter de oorzaken van placebo-respons in depressie- of het stoppen van patiënten in schizofrenie-studies (**Hoofdstuk 2** en **6**). Als er een of meer voorspellende factoren waren gevonden, zouden deze zeker in overweging worden genomen bij de opzet en/of uitvoer van nieuwe studies in de betreffende indicatie(s). Ondanks een uiterst zorgvuldige opzet en uitvoering van het onderzoek in **Hoofdstuk 3**, met innovatieve second opinion-beoordeling van klinische symptomen en geavanceerde statistieken, kon geen significant effect van Org 25935 op negatieve symptomen bij schizofrenie worden vastsgesteld; het kleine effect dat werd waargenomen, bleek hoofdzakelijk niet-specifiek te zijn, d.w.z. secundair aan veranderingen in positieve, algemene psychopathologische, depressieve, cognitieve en extrapiramidale symptomen, en ook daarom onbeduidend. Als een specifiek, statistisch significant effect op negatieve symptomen zou zijn gevonden, is het niet onwaarschijnlijk dat de innovatieve, second-opinion-beoordeling, die voor het eerst in deze studie werd gebruikt, in de toekomst meer algemeen zou worden toegepast om subjectiviteit bij geneesmiddelenonderzoek te verminderen. Vooral, wanneer zou zijn aangetoond dat de effectgroottes toenemen wanneer beoordelaars hun voorlopige scores aanpassen op basis van second opinion. En eventueel zouden de gebruikte analyse-methodes vaker toepassing vinden bij de evaluatie van geneesmiddeleffecten op nauw verweven symptoomcomplexen, wanneer was aangetoond dat Org 25935 daadwerkelijk een specifiek effect had op negatieve symptomen. Maar helaas, met het uitblijven van een gewenst effect zijn de verdiensten van deze innovatieve studieaanpak op langere termijn slechts denkbeeldig.

De resultaten van studies, gepresenteerd in **Hoofdstuk 4, 5** en **7**, tonen aan dat een experimentele behandeling met asenapine en mirtazapine niet inferieur is aan een behandeling met de bestaande geneesmiddelen fluvoxamine en olanzapine, bij respectievelijk schizofrenie en depressie. Tegelijkertijd kon aan de hand van de resultaten geen betekenisvolle verbetering ten opzichte van de gevestigde producten worden aangetoond en leverden de studies geen

wezenlijk nieuwe inzichten op ten aanzien van de behandeling van deze psychiatrische aandoeningen.

Echter, ondanks een gebrek aan potentieel invloedrijke resultaten, verdienen dit soort studies evenveel aandacht als die met een positiever resultaat. Ten eerste: hoewel er duidelijk aanwijzingen zijn dat sommige symptomen bij schizofrenie samenhangen met een ondermaatse, glutamaterge neurotransmissie, is het falen van twee bedrijven om door middel van glycine-opnameremming de symptomen te doen verbeteren een belangrijk waarschuwingssignaal dat aandacht verdient.[8-13] Het delen van dit soort negatieve resultaten door middel van publicatie blijft belangrijk om derden ervoor te behoeden dat zij opnieuw patiënten onnodig blootstellen aan experimenten met vergelijkbare, niet-werkzame stoffen. Ten tweede: Op basis van zowel korte-, als lange-termijn-resultaten, gepresenteerd in **Hoofdstuk 4, 5** en **7**, kunnen zowel asenapine als mirtazapine worden genomen als bruikbare alternatieven voor patiënten die vervelende bijwerkingen ondervinden van voorgeschreven geneesmiddelen.

Het feit dat ik geen voorspelbare oorzaken van placebo-respons in depressie of het voortijdig stoppen van patiënten in een schizofrenie-studie kon identificeren, maakt het in dit stadium niet wenselijk om bepaalde patiënten uit te sluiten van geneesmiddelonderzoek op basis van enquête-uitslagen of gepubliceerde meta-analyses, in een poging om snel een gewenst resultaat boven tafel te krijgen. Evenmin, om het verder speuren naar mogelijkheden voor proefoptimalisatie door middel van meta-analyses als ongepast te beschouwen en maar volledig op te geven. In een recent artikel opperen Cipriani en collega's (2018) goede mogelijkheden voor het combineren van individuele patiëntgegevens uit gerandomiseerde studies in zogenaamde netwerk-meta-analyses.[14] Ze spreken de verwachting uit dat dergelijke analyses beter geschikt zijn dan de conventionele, om de uitkomst van patientgerichte behandeling, de relatieve werkzaamheid van bepaalde middelen, of het risico op vroegtijdige beëindiging te voorspellen. De effectgroottes van antidepressiva en antipsychotica zijn bij klinische studies doorgaans vrij bescheiden, wat volgens menigeen suggereert dat deze geneesmiddelen in hun geheel nauwelijks effectiever zijn dan placebo.[14,15] Er zijn gegevens die erop wijzen, dat nog niet eerder behandelde patiënten vaak beter reageren op een medicijn dan chronische zieke (al eerder behandelde) patiënten met schizofrenie.[16] Deze trend was ook waarneembaar in de naturalistische studie die we hebben uitgevoerd (**Hoofdstuk 4**, gegevens niet getoond). Dit lijkt perspectieven te bieden voor het aangaan van studies waarbij alleen

patiënten, die in hun eerste episode verkeren, mogen deelnemen teneinde de efficiëntie van programma's voor geneesmiddelontwikkeling te verhogen. Daar staat tegenover, dat nieuwe patiënten relatief dun gezaaid zijn en moeilijk te motiveren voor deelname aan een studie, zodat een dergelijke strategie slechts beperkt kan worden toegepast. Kleinschalige studies worden doorgaans tijdens de vroege ontwikkeling van een geneesmiddel uitgevoerd. Het is echter precies dan, wanneer het bewijs van de werkzaamheid nog niet is bewezen en patiënten het moeilijkst zijn te motiveren om zich vrijwillig in te schrijven voor een klinisch onderzoek. Hoewel verbeterde selectie van patiënten in theorie nuttig zou kunnen zijn bij de ontwikkeling van geneesmiddelen, zou het veld in de toekomst toch meer baat kunnen hebben bij een geïntegreerde verwerking van inmiddels verzamelde gegevens, zoals door middel van de hierboven genoemde netwerk-meta-analyse. Open toegang tot de gegevens van zoveel mogelijk (liefst alle) gerandomiseerde studies is cruciaal om deze benadering een kans van slagen te geven. Het vaststellen van responsbevorderende factoren maakt het mogelijk, dat klinische studies optimaal worden ingericht om de effectiviteit van een nieuwe behandeling aan te tonen. Ofschoon dit zeker zal bijdragen tot de repliceerbaarheid van studies en kostenbesparingen, is het maar de vraag of we daarmee tevens in staat zullen zijn om het algehele ontwikkelingstraject van nieuwe geneesmiddelen tegen neuropsychiatrische aandoeningen te verkorten. Dat doel kan gewoonweg te ambitieus zijn, zolang er geen objectieve criteria en / of biomarkers zijn voor de etiologie, diagnose en beoordeling van aandoeningen van het menselijk brein.

REFERENTIES

1. Schoemaker JH. Benefits of placebos. *Nature* 1995;377(6545):98-98.
2. Umbricht D, Alberati D, Martin-Facklam M, Borroni E, Youssef EA, Ostland M, Wallace TL, Knoflach F, Dorflinger E, Wettstein JG, Bausch A, Garibaldi G, Santarelli L. Effect of bitopertin, a glycine reuptake inhibitor, on negative symptoms of schizophrenia: a randomized, double-blind, proof-of-concept study. *JAMA Psychiatry* 2014;71(6):637-646.
3. Bugarski-Kirola D, Iwata N, Sameljak S, Reid C, Blaettler T, Millar L, Marques TR, Garibaldi G, Kapur S. Efficacy and safety of adjunctive bitopertin versus placebo in patients with suboptimally controlled symptoms of schizophrenia treated with antipsychotics: results from three phase 3, randomised, double-blind, parallel-group, placebo-controlled, multicentre studies in the SearchLyte clinical trial programme. *The Lancet Psychiatry* 2016;3(12):1115-1128.
4. Bugarski-Kirola D, Wang A, Abi-Saab D, Blättler T. A phase II/III trial of bitopertin monotherapy compared with placebo in patients with an acute exacerbation of schizophrenia–results from the CandleLyte study. *European Neuropsychopharmacology* 2014;24(7):1024-1036.
5. Siu J. New magic or old tricks? A psychiatry update. *CSHP BC Branch Clinical Symposium, Vancouver (BC)* 2013.
6. Newcomer JW. Antipsychotic medications: metabolic and cardiovascular risk. *Journal of Clinical Psychiatry* 2007;68:8-13.
7. ICH. International Conference on Harmonization Tripartite Guidance E5: Ethnic Factors in the Acceptability of Foreign Data. *The US Federal Register* 1998;63(111):31790–31796.
8. Hons J, Zirko R, Ulrychova M, Cermakova E, Doubek P, Libiger J. Glycine serum level in schizophrenia: relation to negative symptoms. *Psychiatry Research* 2010;176(2):103-108.
9. Sur C, Kinney GG. The therapeutic potential of glycine transporter-1 inhibitors. *Expert Opinion on Investigational Drugs* 2004;13(5):515-521.
10. Gomeza J, Ohno K, Betz H. Glycine transporter isoforms in the mammalian central nervous system: structures, functions and therapeutic promises. *Current Opinion in Drug Discovery & Development* 2003;6(5):675-682.
11. Hutson P, Clark J, Cross A. CNS target identification and validation: avoiding the valley of death or naive optimism? *Annual Review of Pharmacology and Toxicology* 2017;57:171-187.
12. Kantrowitz JT. Managing Negative Symptoms of Schizophrenia: How Far Have We Come? *CNS Drugs* 2017;31(5):373-388.
13. Cioffi CL. Glycine transporter-1 inhibitors: a patent review (2011-2016). *Expert Opinion on Therapeutic Patents* 2018(just-accepted).
14. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, Leucht S, Ruhe HG, Turner EH, Higgins JP. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *The Lancet* 2018.
15. Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B, Samara M, Rabaioli M, Bächer S, Cipriani A. Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, bayesian meta-analysis, and meta-regression of efficacy predictors. *American Journal of Psychiatry* 2017;174(10):927-942.
16. Lieberman J, Jody D, Geisler S, Alvir J, Loebel A, Szymanski S, Woerner M, Borenstein M. Response in First-Episode Schizophrenia. *Archives of General Psychiatry* 1993;50:369-376.

APPENDIX B

List of abbreviations (glossary)

5-HT	Serotonin / 5-hydroxytryptamine (neurotransmitter)
ACTAMESA	Active-Controlled, Two-Armed, Multicenter, Efficacy and Safety Assessment trial (acronym)
ADR	Assessment of Drug Relationship of an AE (safety parameter)
AE	Adverse event (safety parameter)
AIMS	Abnormal Involuntary Movement Scale (rating instrument)
ALT	Alanine Aminotransferase (liver health biomarker in plasma)
ANCOVA	Analysis Of Covariance (statistical method)
ANOVA	Analysis Of Variance (statistical method)
AST	Aspartate Aminotransferase (liver health biomarker in plasma), *or* All-Subjects-Treated (trial participants having received medication)
BARS	Barnes Akathisia Rating Scale (rating instrument)
BID	Twice Daily (Latin: bis in die)
BL	Baseline (cut-off point for evaluation of efficacy and safety)
BMI	Body Mass Index (health measure)
CATIE	Clinical Antipsychotic Trials of Intervention Effectiveness (acronym)
CDSS	Calgary Depression Scale for Schizophrenia (rating instrument)
CFB	Change From Baseline (outcome measure)
CGI	Clinical Global Impression scale (rating instrument)
CGI-I	Clinical Global Improvement score (rating instrument)
CGI-S	Clinical Global Severity score (rating instrument)
CI	Confidence Interval (summary statistics)
CNS	Central Nervous System (brain and spinal cord)
CPT	Continuous Performance Test (cognitive test)
CSF	Cerebro-Spinal Fluid (body fluid)
CYP	Cytochrome P450 (liver metabolizing enzyme)
D (*d*)	Delta value (difference in response between groups)
D_2	Dopamine-2 (neurotransmitter subtype)
D-15	Farnsworth Dichotomous test (color vision test)
DSM-IV	Diagnostic and Statistical Manual, 4^{th} edition (disease criteria)
ECG	Electrocardiogram (electrical heart activity recording)
ECT	Electro-Convulsive Therapy (seizure-inducing shock treatment)
EP	Endpoint (cut-off point for evaluation of efficacy and safety)
EPS	Extra-Pyramidal Symptoms (parkinson's disease mimicking effects)
ERASE	Expert Rater Assisted Score Evaluation (rating control method)
ESRS	Extrapyramidal Symptom Rating Scale (rating instrument)
ETDRS	Early Treatment Diabetic Retinopathy Study (visual acuity test)
FGA	First Generation Antipsychotics (drug class)

GAF	Global Assessment of Functioning (rating instrument)
GCP	Good Clinical Practice (international ethical and scientific quality standard)
GIANT	Glycine uptake Inhibitor Add-on in Negative symptoms Trial (acronym)
GLP	Good Laboratory Practice (quality standard)
GlyT-1	Glycine Transporter Type 1 (co-agonist transporter molecule)
GSR	Gold Standard Rating (rater validation criterion)
γ-GT	Gamma-Glutamyl Transferase (liver health biomarker in plasma)
HAM-D	Hamilton Depression Rating Scale (rating instrument)
IP	Investigational Product (experimental treatment)
ISST	InterSePT Scale for Suicidal Thinking (rating instrument)
IST	Investigator Satisfaction with Treatment (outcome measure)
ITE	Insufficient Therapeutic Effect (outcome measure)
ITT	Intent-To-Treat (trial participants with ≥1 post-baseline assessment)
K_i	Inhibitory Constant (pharmacodynamic parameter)
LFU	Lost to Follow-Up (outcome measure)
LOCF	Last Observation Carried Forward (statistical method)
LOF	Level of Function (rating instrument)
LS	Least Square (summary statistics)
MADRS	Montgomery & Åsberg Depression Rating Scale (rating instrument)
MDD	Major Depressive Disorder (disease)
MMRM	Mixed-Effect Model Repeated Measure (statistical method)
N (*n*)	Number (count)
NCI	Neuro-Cognitive Index (cognitive functioning measure)
NDA	New Drug Application (full dossier for requesting registration of a new drug)
NES	Neurological Evaluation Scale (rating instrument)
NMDA	Glutamate / N-methyl-D-aspartate (neurotransmitter)
NSS	Neurological Soft Signs (rating instrument)
OC	Observed Cases (statistical method)
OQAQ	Overview Quality Assessment Questionnaire (quality measure)
P (*p*)	Probability Value (statistical significance of evidence)
PANSS	Positive and Negative Syndrome Scale (rating instrument)
PANSS-M	PANSS Marder factor score for negative symptoms (symptom subscore)
PET	Positron Emission Tomography (imaging technique)
PK/PD	Pharmacokinetics / Pharmacodynamics (pharmacology)
POET	Perception Of Emotions Test (cognitive test)
PST	Patient Satisfaction with Treatment (outcome measure)
QD	Once daily (Latin: quaque die)
QT_c	Corrected heart frequency QT interval (conduct velocity measure)
RCT	Randomized, Controlled Trial (research method)
R&D	Research & Development (industrial term)
RDC	Research Diagnostic Criteria (disease criteria)
SAD	Schizo-Affective Disorder (DSM-IV differential diagnosis)
SAE	Serious Adverse Event (safety parameter)
SANS	Scale for Assessment of Negative Symptoms (rating instrument)
SAS	Simpson-Angus Scale (rating instrument), *or* Statistical Analysis System (software)
SCZ	Schizophrenia (DSM-IV differential diagnosis)
SD	Standard Deviation (summary statistics)

SE	Standard Error (summary statistics)
SF-12	12-Item Short Form health survey (rating instrument)
SGA	Second Generation Antipsychotics (drug class)
SGOT	Serum Glutamic Oxaloacetic Transaminase = AST
SGPT	Serum Glutamate-Pyruvate Transaminase = ALT
SOHO	Schizophrenia Outpatient Health Outcomes trial (acronym)
SPSS	Statistical Package for the Social Sciences (software)
SWN	Subjective Well-being under Neuroleptic treatment scale (rating instrument)
TCA	Tricyclic antidepressant (drug)
WHO	World Health Organization (specialized agency of the United Nations)
WOC	Withdrawal Of Consent (outcome measure)

APPENDIX C
List of publications

JOURNAL ARTICLES

Schoemaker JH, Kilian, S, Emsley R., Vingerhoets, A. Factors associated with placebo response in depression trials: A systematic review of published meta-analyses (1990-2017). *Neurology, Psychiatry, and Brain Research* 2018;30:12-21.

Schoemaker JH, Vingerhoets A, Emsley R. Factors associated with poor satisfaction with treatment and trial discontinuation in chronic schizophrenia. *CNS Spectrums* 2018; doi:10.1017/S109285291700044X.

Bejczy A de, Nations K, Szegedi A, **Schoemaker J**, Ruwe F, Söderpalm B. Efficacy and safety of Org 25935 for the prevention of relapse in alcohol-dependent patients. *Alcoholism: Clinical and Experimental Research* 2014;38(9):2427-2435.

Schoemaker JH, Jansen WT, Schipper J, Szegedi A. The selective glycine uptake inhibitor Org 25935 as an adjunctive treatment to atypical antipsychotics in predominant persistent negative symptoms of schizophrenia: results from the GIANT trial. *Journal of Clinical Psychopharmacology* 2014;34(2):190-198.

Schoemaker J, Stet L, Vrijland P, Naber D, Panagides J, Emsley R. Long-term efficacy and safety of asenapine or olanzapine in patients with schizophrenia or schizoaffective disorder: an extension study. *Pharmacopsychiatry* 2012;45(05):196-203.

Schoemaker J, Naber D, Vrijland P, Panagides J, Emsley R. Long-term assessment of asenapine vs. olanzapine in patients with schizophrenia or schizoaffective disorder. *Pharmacopsychiatry* 2011;44(07):343.

De Greef R, Malony A, Olsson-Gislekog P, **Schoemaker J**, Panagides J. Dopamine D2 occupancy as a biomarker for antipsychotics: quantifying the relationship with efficacy and extrapyramidal symptoms. *American Association of Pharmaceutical Scientists Journal* 2011;13(1):121-130.

Schoemaker J, Naber D, Vrijland P, Panagides J, Emsley R. Long-term assessment of asenapine vs. Olanzapine in patients with schizophrenia or schizoaffective disorder. *Pharmacopsychiatry* 2010;43(04):138-146.

Murasaki M, **Schoemaker JH**, Miyake M, Gailledreau J, Heukels AJ, Fennema HP, Sitsen JMA. Comparison of efficacy and safety of mirtazapine versus fluvoxamine in Japanese and Caucasian patients with major depressive disorder. *Japanese Journal of Clinical Psychopharmacology* 2010;13:339-355.

Droogleever Fortuyn HA, **Schoemaker JH**. Treatment of delirium with phototherapy. *European Psychiatry* 1997;12(7):367-368.

Schoemaker JH. Benefits of placebo. *Nature* 1995;377:98.

Schoemaker JH, Bousema MT, Zijlstra H, Van der Horst FAL. Effective treatment of erythropoietic protoporphyria with hydroxyethylrutosides. *Dermatology* 1995;191:36-38.

Schoemaker JH. Impaired axonal transport in diabetic neuropathy. *Diabetes Care* 1994;17(11):1362.

Schoemaker JH. Pharmacological treatment of diabetic peripheral neuropathy. *British Journal of Clinical Practice* 1994;48(2):91-96.

PUBLISHED ABSTRACTS & POSTERS

Zhao J, Cazorla P, **Schoemaker J**, Mackle M, Panagides J, C Karson, A Szegedi. Weight change and metabolic effects of asenapine in placebo- or olanzapine-controlled studies. *European Psychiatry* 2011;26:250.

Schoemaker J, Naber D, Jansen W, Panagides J, Emsley R. Safety and efficacy of long-term asenapine versus olanzapine in schizophrenia or schizoaffective disorder patients. *European Psychiatry* 2010;25:1621-1621.

Schoemaker J, Stet L, Naber D, Panagides J, Emsley R. Long-term efficacy and safety of asenapine versus olanzapine in patients with schizophrenia or schizoaffective disorder: an extension study. *163rd Annual Meeting of the American Psychiatric Association (APA)*, May 22-26 May 2010, New Orleans, LA.

Gaur R, Ramirez L, De Santi S, **Schoemaker J**. Rater training on SANS and monitoring of rater performance during clinical trials. *European Neuropsychopharmacology* 2009;1:19.

Gaur R, Sajatovic M, **Schoemaker J**, DeSanti S. A survey of rater perceptions and expectations of a rater training program in a CNS drug trial. *15th Biennial Winter Workshop in Psychoses,* 15-18 November 2009, Barcelona.

Schoemaker J, Gaur R, Sajatovic M, Ramirez L, De Santi S. An interview guide for the scale for assessment of negative symptoms (IG-SANS). *European Archives of Psychiatry and Clinical Neuroscience* 2009;259 (Suppl 1):67.

Gaur R, Sajatovic M, Rijhwani M, **Schoemaker J**, De Santi S. Signal enhancement system (SES) identifies inaccurate ratings and quality of clinical interview conducted in a CNS trial. *22nd ECNP Congress*, 12-16 September 2009, Istanbul, Turkey.

Schoemaker J, Gaur R, Chawla V, Szegedi A. Expert rater assisted score evaluation: a new method to increase efficiency in drug development programs. *22nd ECNP Congress*, 12-16 September 2009, Istanbul, Turkey.

De Greef R, Malony A, Olsson-Gislekog P, Prins K, **Schoemaker J**, Panagides J. Application of mathematical models based on D2 occupancy. *New Clinical Drug Evaluation Unit (NCDEU) 49th Annual Meeting*, 29 June – 02 July 2009, Hollywood, Florida.

De Greef R, Malony A, Olsson-Gislekog P, Prins K, **Schoemaker J**, Panagides J. Dose selection of asenapine: application of models based on D2 occupancy to predict efficacy and EPS. *American Psychiatric Association 162nd Annual Meeting*, 16-21 May, 2009, San Francisco, CA.

Schoemaker JH, Szegedi A. Factors enhancing placebo response in depression trials. *International Society for CNS Clinical Trials & Methodology (ISCTM) Mid-Year Conference*, 17-18 September 2007, Brussels.

Kleijn HJ, Gerrits MGF, **Schoemaker JH**, De Greef HJMM. PK-PD modeling and simulation of phase I data to support development of a new compound in the field of CNS. *5th International Symposium on Measurement and Kinetics of In Vivo Drug Effects*, 26-29 April 2006, Noordwijkerhout, The Netherlands.

Schoemaker J, Gailledreau J, Høyberg OJ. First, randomized, double-blind comparison of mirtazapine (15-45 mg) and fluvoxamine (50-150 mg) in the treatment of depression. *International Journal of Neuropsychopharmacology* 2002;5(Suppl 1):140.

Schoemaker JH, Murasaki M, Miyake M. Randomized, double-blind comparison of mirtazapine (15-45 mg) and fluvoxamine (50-150 mg) in the treatment of depression in Japan. *XII World Congress of Psychiatry*, 24-29 August 2002, Yokohama.

ORAL PRESENTATIONS AT INTERNATIONAL CONFERENCES

Schoemaker JH. A novel web-based technique to confirm diagnosis and enhance signal detection in randomized controlled trials (oral presentation). *Dermatological Advances*, 04 December, 2009. London.

Schoemaker JH. Signal enhancement in psychiatric patients. *5th Latin American Congress of Clinical Research (DIA)*, 19-21 November, 2008, Buenos Aires.

APPENDIX D

Acknowledgements (dankwoord)

Veel dank ben ik verschuldigd aan mijn beide promotores, Ad Vingerhoets en Robin Emsley. Ad, zonder jouw suggestie om te promoveren was het er wellicht nooit van gekomen. Je geduldige begeleiding en oog voor detail hebben mij veel geholpen om de teksten leesbaar en toegankelijk te houden, hopelijk ook voor niet-ingewijden. Robin, I am greatly indebted to you for your expertise in the field, kind collaboration over many years, and encouragement to proceed with my study approaches despite all obstacles and politics, which showed up now and then.

Rinus Verkooijen heeft in belangrijke mate bijgedragen tot het voltooien van de manuscripten in de gewenste stijl. Rinus, dankzij jouw toegewijding en doorzettingsvermogen kon de tekstopmaak van eerder gepubliceerde studies vlot worden omgebogen naar de vorm die wenselijk was. Ik ben je ontzettend dankbaar voor je inzet, die mij geweldig veel tijd heeft bespaard.

Bij aanvang van mijn post-hoc analyses heb ik veel steun gehad aan Wobbe Zijlstra, die zo aardig was om tussen zijn drukke werkzaamheden door, enkele data voor mij te bewerken in R. Wobbe, alsnog reuzebedankt hiervoor.

Zowel voor als tijdens de uitvoering van mijn analyses mocht ik mij verheugen in de warme belangstelling van velen binnnen de vakgroep Medical and Clinical Psychology, waarvoor mijn dank. Ofschoon werkzaam als buiten-promovendus, voelde ik mij daardoor opgenomen als collega binnen de vakgroep. Niet op de laatste plaats ook dank aan Saskia Diesbergen die erop toezag dat ik steeds over een vaste werkplek kon beschikken.

De studies in hoofdstukken 3, 4, 5 en 7 kwamen tot stand dankzij intensieve samenwerking met vele collega's van diverse afdelingen binnen Organon (later Schering-Plough, respectievelijk Merck, Sharpe & Dohme). Ofschoon ik zelf altijd aan het roer heb gestaan, zowel bij de opzet en uitvoer van de studies als bij de uitwerking, rapportage en publicatie van resultaten, zouden deze studies nooit mogelijk zijn geweest zonder onophoudelijke samenwerking van vele specialisten binnen de diverse onderzoeksteams. Dankzij (maar soms ook ondanks) de expertise van velen was ik in staat om de projecten binnen de daarvoor gestelde tijdslijnen uit te voeren en af te ronden. Twee personen zou ik hier met naam willen noemen, zonder wiens aanwezigheid ik daar nooit in zou hebben kunnen slagen. Op de eerste plaats Ad Sitsen, groepshoofd van de sectie Psychiatrie aan de afdeling Clinical Drug Development bij Organon in de tijd dat ik daar solliceerde naar de functie van Senior Medical Research Scientist. Ad, als jij me niet had aangenomen weet ik niet hoe mijn carrière verder was gelopen aangezien het aantal bedrijven in Nederland dat geneesmiddelen ontwikkelt altijd heel schaars is geweest. Ik heb mij lange tijd mogen verheugen in een enthousiaste samenwerking met jou en ben je nog altijd zeer dankbaar voor het vertrouwen dat je in mij hebt gesteld. Op de tweede plaats Wim Jansen, statisticus aan de afdeling Biometrie. In tegenstelling tot menig collega van Wim, was hij altijd bereid om alternatieve benaderingen te bespreken en additionele analyses uit te voeren als deze iets wezenlijks konden toevoegen.

Wim, als jij je nek niet had durven uitsteken had het resultaat van de studie in hoofdstuk 3 beslist minder eenduidig geweest.

Ook een woord van dank aan mijn vrouw Freke, van wie ik onvoorwaardelijk alle tijd en gelegenheid heb gekregen om mijn studies uit te voeren en dit proefschrift te voltooien. Freke, lieve schat, jouw vertrouwen in mij heeft mij steeds voortgestuwd, en doet dit nog altijd (meer dan je misschien beseft).

Tot slot wil ik benadrukken dat ik hier zonder de liefderijke opvoeding van mijn ouders, die mij altijd de ruimte hebben gegeven om mijzelf te ontplooien in de richting die ik zelf verkoos, nooit zou hebben gestaan. Mogen zij rusten in vrede.

APPENDIX E
Curriculum vitae

Joep Schoemaker werd op 11 oktober 1958 geboren in Vught. Na het behalen van het Gymnasium β diploma is hij in 1977 Biologie gaan studeren aan de Faculteit Wiskunde en Natuurwetenschappen van de Universiteit Utrecht, en sloot deze studie met een doctoraalexamen af in 1984. Van 1985 tot en met 2014 is Joep werkzaam geweest in diverse functies binnen de farmaceutische industrie. Zo werkte hij bij grote multinationals als Hoffman – La Roche (1985-1997), N.V. Organon (1997-2007), Schering-Plough (2007-2009) en Merck, Sharpe & Dohme (2009-2011). Gedurende meer dan 25 jaar heeft hij daarbij verantwoordelijkheid gedragen voor de opzet, uitvoer, resultaatverwerking en rapportage van klinische onderzoeksprojecten met nieuwe, nog te registreren geneesmiddelen, in binnen- en buitenland. Aanvullend was Joep als executive director werkzaam van 2011 tot 2015 bij inVentiv Health (thans Syneos Health), een internationale Contract Research Organization, waar hij leiding gaf aan Europese project managers en directors, die waren belast met geneesmiddelenonderzoek in psychiatrische en neurologische aandoeningen.

Reeds tijdens de doctoraalfase van zijn studie Biologie was Joep gefascineerd door de werking van het brein en de mogelijke invloed van endogene, biologische factoren op emotie en gedrag. Ofschoon hij het altijd jammer heeft gevonden dat er in de jaren tachtig toen hij afstudeerde weinig of geen fondsen beschikbaar waren voor fundamenteel onderzoek, en hij geen goede kansen had om zich als wetenschapper verder te verdiepen in de neurofysiologie, heeft hij zich vanaf dag één binnen de industrie met grote interesse en energie beziggehouden met de uitdagingen binnen psychiatrisch en neurologisch onderzoek. Al snel leerde hij dat – met het ontbreken van adequate ziektemodellen en mogelijkheden voor het doen van directe metingen – ieder onderzoek binnen deze gebieden een speciale benadering vereist, waarbij niet alleen gestandaardiseerde patiënteninterviews maar ook therapietrouw en een doordachte onderzoeksopzet vaak aan de basis staan van een geslaagde studie. Geconfronteerd met voortdurend nieuwe uitdagingen, heeft Joep door de jaren heen diverse onderzoeksstrategieën uitgedacht en/of uitgeprobeerd, met wisselend succes. Bundeling van deze strategieën heeft uiteindelijk aan de basis gestaan van dit proefschrift. Onder supervisie van zijn promotoren heeft hij in aanvulling hierop een tweetal generieke factoren die vaak afbreuk doen aan studies binnen de psychiatrie nader onderzocht, te weten: placebo respons en voortijdige studie-uitval. In de hoop dat eventuele, gunstige resultaten hiervan mede zouden kunnen bijdragen tot een zo efficiënt mogelijk maken van toekomstig geneesmiddelenonderzoek in neuro-psychiatrische aandoeningen.